International School for Advanced Studies

Ph.D. Thesis

Quantum annealing and digital quantum ground state preparation algorithms

Author:
Glen Bigan Mbeng

Supervisor:
Prof. Giuseppe E. Santoro

October 21, 2019

"Now, for the first time in human history, we are acquiring and perfecting the tools to build and precisely control very complex, highly entangled quantum states of many particles, states so complex that we can't simulate them with our best digital computers or characterize them well using existing theoretical tools. This emerging capability will open the door to new discoveries."

[Preskill, 2018]

Abstract

The recent advances in the world of quantum technologies have prompted the development of various quantum-based algorithms, some of which are suitable to run on available quantum devices. Two leading candidates in this area are Quantum Annealing (QA) and hybrid quantum-classical variational algorithms, such as the Quantum Approximate Optimization Algorithm (QAOA).

In the first part of the Thesis we re-examine the issue of whether Simulated Quantum Annealing (SQA) with Path Integral Monte Carlo has anything to do with the coherent QA Schrödinger dynamics. We do that by studying the random quantum Ising Chain in a transverse field.

In the second part of the Thesis we address the issue of schedule optimization in both *digitized*-QA and QAOA. Traditional schedule optimization in QA requires spectral information on the problem (e.g., location of the minimum gaps), which is in most cases inaccessible. On the other hand, in QAOA, the issue of schedule optimization — *alias*, the optimization of the variational parameters — is often an expensive task that constitutes a bottleneck for the algorithm. We show that, by combining the framework of QAOA and digitized-QA, optimal digitized-QA protocols can be constructed efficiently. When using this approach, the computational cost of schedule optimization is also significantly reduced, leading to a computational advantage both over a linear-schedule QA, and over an unstructured QAOA search. While studying these issues, we also developed rigorous variational bounds that provide insight on the best possible performance that one can hope to obtain from a digital QAOA circuit.

Acknowledgments

Firstly, I would like to thank my advisor Prof. Giuseppe Santoro for the continuous support of my Ph.D. study. During these four years, his guidance helped me in all the aspects of research, and in writing this Thesis. I especially thank him for the exceptional patience and dedication displayed in this last year.

I thank my family, to whom I dedicate this Thesis. I could not imagine getting anywhere without their support. Dad, Mum, Tina, and Vivy, thank you for helping me overcoming the difficulties and celebrating the successes encountered in these four years of Ph.D.

I thank the various professors, researchers, and fellows Ph.D. students for the stimulating discussions. I especially thank Prof. Rosario Fazio, who also supported me and actively contributed to developing various ideas contained in this Thesis.

Finally, I thank all my friends. They made my stay in Trieste a wonderful experience.

Contents

List of Figures

Publications

The following is a list of the papers, preprints and manuscripts in preparation which will be discussed in the Thesis.

1. G.B. Mbeng, L. Privitera, L. Arceci, and G.E. Santoro, *Dynamics of simulated quantum annealing in random Ising chains*, Phys. Rev. B **99**, 064201 (2019). (Discussed in Chap. 2.)

2. G.B. Mbeng, L. Arceci, and G.E. Santoro, *Optimal working point in digitized quantum annealing*, arXiv:1909.00817 (2019). Sumbitted to Physical Review B. (Discussed in Chap. 3.)

3. G.B. Mbeng, R. Fazio, and G.E. Santoro, *Quantum Annealing: a journey through Digitalization, Control, and hybrid Quantum Variational schemes*, arXiv:1906.08948 (2019). A modified version of this manuscript will be soon submitted to Physical Review B. (Discussed in Chap. 4.)

4. G.B. Mbeng, R. Fazio, and G.E. Santoro, *Optimal quantum control with digitized quantum annealing*, manuscript in preparation. (Discussed in Chap. 4.)

5. G.B. Mbeng, R. Fazio, and G.E. Santoro, *Bounds on quantum control for ground-state preparation*, manuscript in preparation. (Discussed in Chap. 5.)

6. G.B. Mbeng, R. Fazio, and G.E. Santoro, *Symmetry breaking, speed limits and bounds on quantum variational schemes for ground-state preparation*, manuscript in preparation. (Discussed in Chap. 5.)

Below is a list of other works and projects, performed during my PhD, which **will not be discussed in the Thesis**.

1. G.B. Mbeng, V. Alba, and P. Calabrese, *Negativity spectrum in 1D gapped phases of matter*, J. Phys. A: Math. Theor. **50**, 194001 (2017).

2. A. Achille, G. Mbeng, and S. Soatto, *The Dynamics of Differential Learning I: Information-Dynamics and Task Reachability*, arXiv:1810.02440 (2018).

3. A. Achille, G. Paolini, G. Mbeng, and S. Soatto, *The Information Complexity of Learning Tasks, their Structure and their Distance*, arXiv:1904.03292 (2019).

4. M. Wauters, G.B. Mbeng, and G.E. Santoro, *Hybrid quantum-classical variational schemes for the fully-connected* p*-spin Ising ferromagnet*, manuscript in preparation.

5. J. Hašík and G.B. Mbeng, *A differentiable tensor network library for two-dimensional lattice models*, soon available on GitHub.

Abbreviations

ABC	**A**ntiperiodic **B**oundary **C**onditions
dQA	**d**igitized **Q**uantum **A**nnealing
KZ	**K**ibble **Z**urek
MCS	**M**onte **C**arlo **S**weep
NP	**N**ondeterministic **P**olynomial time
PIMC	**P**ath **I**integral **M**onte **C**arlo
PBC	**P**eriodic **B**oundary **C**onditions
QA	**Q**uantum **A**nnealing
QAOA	**Q**uantum **A**pproximate **O**ptimization **A**lgorithm
SA	**S**imulated **A**nnealing
SQA	**S**imulated **Q**uantum **A**nnealing

Dedicated to my parents

Chapter 1

Introduction

1.1 Ground state preparation: the problem

Quantum Mechanics is the fundamental theory that describes Nature at the atomic scale [Sakurai and Commins, 1995]. As such, various branches of science and engineering aiming at studying, controlling or manipulating molecular [Helgaker *et al.*, 2014] or condensed matter systems [Grosso and Parravicini, 2000] are based on a quantum mechanical description. In particular, when attempting to give a low-energy (or temperature) description of a complex system, our current inability to efficiently compute the ground state of interesting complex Hamiltonians is a severe bottleneck and hindrance to progress in these fields. Exploring methods to overcome this bottleneck is the central topic of this Thesis.

In general terms, we formulate the "Hamiltonian-to-ground-state problem" as follows. Given a physical system described by a Hamiltonian $\widehat{H}_{\mathrm{T}}$, extract its low-energy proprieties as encoded in its ground state $|\psi_{\mathrm{gs}}\rangle$, see sketch in Fig. 1.1. Our understanding of Nature would tremendously advance if such a device was available. Indeed, many problems in Condensed Matter are still open and highly debated: for instance, the mechanism behind high-Tc superconductivity [Keimer *et al.*, 2015] Equally abundant and relevant, in a more application-oriented setting of quantum chemistry, are the problems connected with the design of innovative drugs or materials [Weymuth and Reiher, 2014].

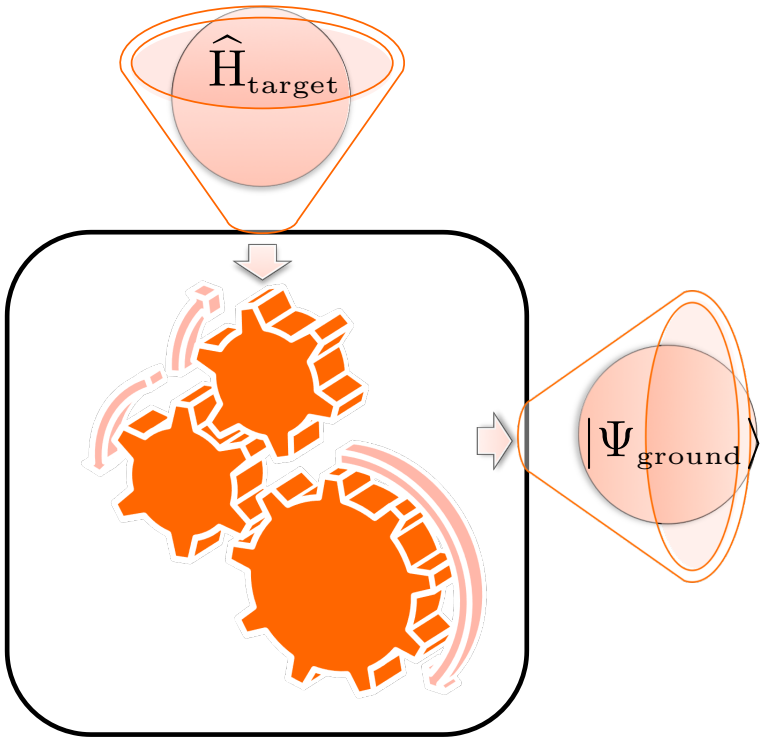

FIGURE 1.1: Illustration of the "H-to-psi-machine": it takes as input a target Hamiltonian $\widehat{\mathrm{H}}_{\mathrm{target}}$ (abbreviated to $\widehat{H}_{\mathrm{T}}$ in the main text) and returns an approximation to it's ground state $|\Psi_{\mathrm{ground}}\rangle$ (abbreviated to $|\psi_{\mathrm{gs}}\rangle$ in the main text).

The most challenging quantum many-body problems[1], indeed, escape[2] our current ability of numerically simulating them with the most sophisticated Quantum Monte Carlo techniques [Becca and Sorella, 2017].

Moreover, even at the level of classical physics, there are many computer science optimization problems (such as Boolean Satisfiability, Travelling Salesman, ...) which are hard — the so-called NP-hard problems — and for which a hypothetical device as that illustrated in Fig. 1.1 would be useful [Boros and Hammer, 2002]. Indeed, most optimization problems can be formulated as the problem of finding the ground state of a classical Ising-like Hamiltonian[3] [Lucas, 2014].

[1]The bottleneck present in all quantum mechanical many-body problems is the exponential growth of the Hilbert space dimension, which one can be appreciated even for a quantum spin system: describing N interacting spin-$\frac{1}{2}$ objects requires 2^N complex variables, which makes an exact description impossible on a classical computer when N exceeds $40-50$.

[2]We recall the infamous *sign-problem* which plagues Quantum Monte Carlo studies of fermionic or frustrated bosonic problems [Troyer and Wiese, 2005].

[3]Once again, the exponential increase of the configuration space, 2^N for N classical Ising variables, is the crucial ingredient that makes these problems "hard".

As proposed by Feynman [Feynman, 1982], some of the quantum problems mentioned above might become much easier for a *quantum computer*: since Nature is quantum, it should be more naturally "simulated" on a device implementing the laws of Quantum Mechanics. [4] Indeed, intense research is currently underway in the field of *quantum simulators*, i.e., laboratory realizations— often with cold atomic gases trapped in optical lattices — which emulate some of the most interesting models of strongly correlated quantum systems [Bloch *et al.*, 2012; Bernien *et al.*, 2017]. More generally, Feynman's ideas have created new fields of research, which go under the name of *Quantum Computation and Quantum Information* [Nielsen and Chuang, 2000]. While the experimental realizations of such "quantum hardware" is still in its infancy, the development of various quantum-based algorithms, some of which are suitable to run on currently available quantum devices — broadly belonging to the class of Noisy Intermediate-Scale Quantum (NISQ) technologies [Preskill, 2018] — is actively pursued. Two leading candidates in this area are Quantum Annealing, and hybrid quantum-classical variational algorithms.

Quantum Annealing (QA) [Kadowaki and Nishimori, 1998; Brooke *et al.*, 1999; Santoro *et al.*, 2002] [5] — *alias* Adiabatic Quantum Computation [Farhi *et al.*, 2001; Albash and Lidar, 2018] — is a strategy, based on the adiabatic theorem [Messiah, 1962] of Quantum Mechanics, for solving interesting hard problems through a continuous-time adiabatic evolution of an appropriate quantum Hamiltonian.

Hybrid quantum-classical variational algorithms, instead, are based on classical minimization and invoke a quantum digital processor to prepare a variational *Ansatz* for the problem [Farhi *et al.*, 2014; Peruzzo *et al.*, 2014; Kokail *et al.*, 2019].

We now discuss these algorithms and related issues in more detail.

1.2 Quantum Annealing

According to standard continuous-time QA [Kadowaki and Nishimori, 1998; Santoro *et al.*, 2002; Farhi *et al.*, 2001; Albash and Lidar, 2018], the target Hamiltonian $\widehat{H}_{\mathrm{T}}$ has to be supplemented by a driving term $\widehat{H}_{\mathrm{D}}$. In the simplest setting, one

[4] This is not to imply that all hard problems of physics and computer science are expected to be "made simple" by a quantum mechanical approach. Indeed, glasses have a slow dynamics. Hence they are in some sense "hard" even for Nature.

[5] See also [Finnila *et al.*, 1994] for an earlier formulation of the concept of optimization through quantum mechanics.

would then write a QA Hamiltonian of the form:

$$\widehat{H}(s) = s\,\widehat{H}_{\mathrm{T}} + (1-s)\,\widehat{H}_{\mathrm{D}}\ . \tag{1.1}$$

The parameter s is then varied in time, defining a schedule $s(t)$ interpolating between $s(0) = 0$ and $s(\tau) = 1$, where τ is the total annealing time. In its simplest form, QA is often associated to a *linear schedule* $s(t) = t/\tau$, but this restriction can be in principle removed, as one might optimise the schedule $s(t)$ appropriately, see Sec. 1.3. Given any schedule $s(t)$, and starting from the ground state $|\psi_0\rangle$ of $\widehat{H}_{\mathrm{D}}$ — assuming this is simple enough to construct —, the state of the system at time t is given by the Schrödinger equation

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \widehat{H}(s(t))|\psi(t)\rangle\ , \tag{1.2}$$

which is formally solved by $|\psi(t)\rangle = \widehat{U}_{\mathrm{QA}}(t,0)|\psi_0\rangle$ where the evolution operator

$$\widehat{U}_{\mathrm{QA}}(t,0) = \mathrm{Texp}\left(-\frac{i}{\hbar}\int_0^t \mathrm{d}t'\,\widehat{H}(s(t'))\right) \tag{1.3}$$

is the time-ordered exponential of the Hamiltonian. The QA algorithm terminates by returning the output state $|\psi(\tau)\rangle$.

As said, the basic idea behind QA is the adiabatic theorem of Quantum Mechanics [Albash and Lidar, 2018]. Indeed, for any finite system, provided the Hamiltonian $\widehat{H}(s)$ has a finite spectral gap $\Delta > 0$ separating its ground state from the higher excited states, the Schrödinger dynamics is guaranteed to lead to the correct target ground state of $\widehat{H}_{\mathrm{T}}$, provided one anneals the system sufficiently slowly, $\tau \to \infty$. As such, QA realises a possible scheme for the "H-to-psi" machine illustrated in Fig. 1.1.

1.2.1 QA for optimization problems

We briefly illustrate the QA approach to solve classical optimization problems through two prototypical examples, one from physics and the other from computer science: the Spin Glass and the MaxCut ground-state problem.

The Spin Glass (SG) ground-state problem is the most natural example of a classical optimization problem arising from physics [Parisi *et al.*, 1987]. A SG, such as the Edwards-Anderson model, is described by a set of N spins (binary variables $S_i = \pm 1$) with a two-body spin-spin interaction. For the Edwards-Anderson case,

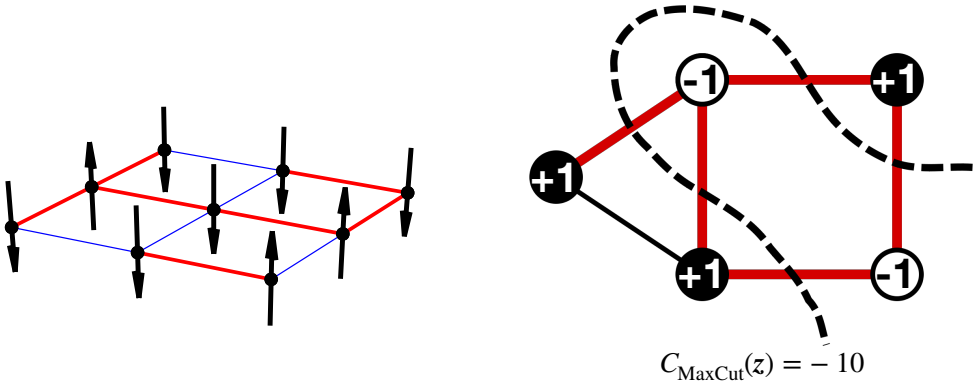

FIGURE 1.2: Illustration of the Spin Glass and the MaxCut ground state preparation problems. (a) Lowest energy configuration of a spin glass problem for $N = 9$ spins. In this instance, the coupling constants J_{ij} are allowed to assume only the values $+1$ (blue lines) and -1 (thick red lines). (b) Bipartition producing the maximum cut (dashed line) for a graph of $N = 5$ nodes. The nodes in the two partitions are all respectively black $(+1)$ or white (-1). The thick red lines represent the cut edges.

the classical Hamiltonian is:

$$E_{SG}(S_1, S_2, \cdots) = -\sum_{\langle ij \rangle} J_{ij} S_i S_j \ , \tag{1.4}$$

where $\sum_{\langle ij \rangle}$ denotes a sum over nearest-neighbor spins on a given lattice and $\{J_{ij}\}$ are real coupling constants. The SG ground-state problem consists in finding a spin configuration that minimizes the energy E_{SG} for a given set of coupling constants $\{J_{ij}\}$, see Fig. 1.2(a). The standard strategy in quantum optimization algorithms is to map $S_j \to \hat{\sigma}_j^z$ and to regard the cost-function as a quantum Hamiltonian for N spin-$\frac{1}{2}$:

$$E_{SG}(S_1, \cdots, S_N) \ \ \to \ \ \widehat{H}_{\text{T}} = E_{SG}(\hat{\sigma}_1^z, \cdots, \hat{\sigma}_N^z) = -\sum_{i<j} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z \ , \tag{1.5}$$

where $\hat{\boldsymbol{\sigma}}_i = (\hat{\sigma}_i^x, \hat{\sigma}_i^y, \hat{\sigma}_i^z)$ stand for the Pauli Matrices representing the i-th spin. In the simplest version of QA, $\widehat{H}_{\text{T}}$ is to be supplemented by a transverse field driving term[6]

$$\widehat{H}_{\text{D}} = -\Gamma \sum_{j=1}^{N} \hat{\sigma}_j^x \ . \tag{1.6}$$

[6] The simple transverse field driving Hamiltonian given in Eq. (1.6) is not necessarily the best choice for QA. For instance, the effect of adding transverse interactions or inhomogeneities to $\widehat{H}_{\text{D}}$ have been studied in [Seoane and Nishimori, 2012] and [Susa *et al.*, 2018].

The starting state is assumed, as usual, to be the ground state of the driving Hamiltonian

$$|\psi_0\rangle = 2^{-N/2}(|\uparrow\rangle + |\downarrow\rangle)^{\otimes N} = |+\rangle^{\otimes N} \ . \tag{1.7}$$

where $|\uparrow\rangle$ and $|\downarrow\rangle$ denote the eigenstates of $\hat{\sigma}^z$, while $|+\rangle$ denotes the eigenstate of $\hat{\sigma}^x$ with eigenvalue $+1$. We then use $\widehat{H}_{\rm D}$ and $\widehat{H}_{\rm T}$ to run the QA algorithm described in the previous section (Eq. (1.1) and Eq. (1.3)).

In the field of computer science, MaxCut is one of the various hard optimization problems that interest the community [Boros and Hammer, 1991]. Given a graph, i.e., a set of N vertices or nodes $\mathcal{G} = \{j\}$ connected by certain edges $\mathcal{E} = \{e\}$, the MaxCut problem consists in finding the largest number of edges that need to be cut when partitioning the graph into two independent parts, as illustrated in Fig 1.2(b). By assigning a label $S_j = +1$ and $S_j = -1$ to the nodes of the two independent parts, the MaxCut problem can be interpreted as the minimization of the cost function

$$C_{\rm MaxCut}(\boldsymbol{z}) = \sum_{\langle i,j\rangle\in\mathcal{E}} (S_i S_j - 1) = -2 \cdot \#(\text{cut edges}) \ , \tag{1.8}$$

where the sum runs on all the edges $\langle i, j\rangle = e \in \mathcal{E}$. This amounts, in the quantum language, to searching for the minimum of an Ising antiferromagnet defined on the graph.

$$\widehat{H}_{\rm T} = J \sum_{\langle i,j\rangle\in\mathcal{E}} (\hat{\sigma}_i^z \hat{\sigma}_j^z - 1) \ , \tag{1.9}$$

where $J > 0$ is an energy scale. As for the SG problem, $\widehat{H}_{\rm T}$ and the transverse field $\widehat{H}_{\rm D}$ in Eq. (1.6) can be used to run the QA algorithm.

1.2.2 Issues with QA: the small spectral gaps

In the QA framework, the hardness of a problem is associated with the intrinsic difficulty in following the adiabatic ground state when the system is driven across a quantum critical point [Zurek *et al.*, 2005]. This is especially severe in cases where an exponentially small spectral gap must be crossed to go from the initial state to the final ground state one is searching for [Knysh, 2016; Bapst *et al.*, 2013]. Different strategies have been proposed to cope with such a problem, including heuristic guesses for the initial state [Perdomo-Ortiz *et al.*, 2011], strategies for increasing the minimum gap [Zeng *et al.*, 2016; Zhuang, 2014] or avoiding first-order

lines [Seoane and Nishimori, 2012], and the quantum adiabatic brachistochrone [Rezakhani *et al.*, 2009].

A second important, and related, issue of QA is the choice of the driving schedule, which we are now going to discuss.

1.3 The schedule problem

In the Grover case [Grover, 1997], Roland & Cerf [Roland and Cerf, 2002] have shown that a schedule optimisation is crucial for obtaining a quantum speed-up through QA. Schedule optimization, however, is believed to require, *in general*, spectral information on the problem at hand — where the critical point is located, and what is the spectral gap that the adiabatic evolution has to cope with — and this poses, clearly, a very hard problem [Ambainis, 2013; Cubitt *et al.*, 2015].

Let us discuss this problem in the context of the Grover problem in more detail.

1.3.1 Quantum adiabatic algorithm for Grover's problem

One of the few problems which are provably known to exhibit a quantum speed-up is Grover's unstructured search [Grover, 1997]. In Grover's problem, we are given a database of N unsorted items labeled from 1 to N. The database contains a single marked item $\overline{m}$, which can be recognized when inspected with a given detector (oracle). The task is then to find $\overline{m}$ using the oracle as few times as possible[7].

On a classical hardware, the most efficient solution is to examine the items in any order, until we find the marked one. This requires an average of $N/2$ queries[8]. Grover showed that one can obtain a quadratic speed-up by running the search on a quantum digital hardware that requires only $\mathcal{O}(\sqrt{N})$ queries of the oracle. Later, this was proven to be the best performance achievable with a quantum device [Bennett *et al.*, 1997; Farhi and Gutmann, 1998; Zalka, 1999].

In the QA framework, Grover's problem can be formulated as the search of a ground state of the target (oracle) Hamiltonian

$$\widehat{H}_{\mathrm{T}} = J\left[1 - |\overline{m}\rangle\langle\overline{m}|\right] , \tag{1.10}$$

[7]For instance, the database could be the set of all possible configurations of a classical system of n spins (i.e., binary variables $S_i = \pm 1$). In this case $N = 2^n$ and the string one is searching for is, e.g., $\overline{m} = (+1, -1, +1, +1, -1, \cdots)$.

[8]See also [Hen, 2019] for a discussion on the performance of classical analog algorithms on Grover's problem.

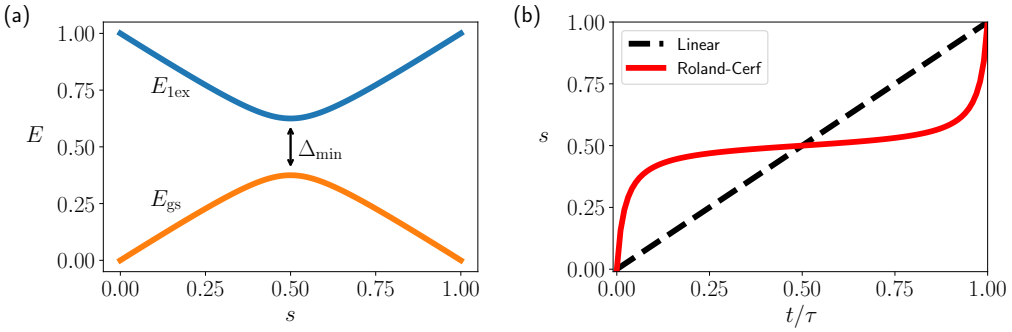

FIGURE 1.3: QA annealing solution to Grover's problem for a database of $N = 32$ items. (a) Energy levels of the ground state $E_{\rm gs}$ and of the first exited state $E_{\rm ex}$ as a function of the interpolation parameter s. (b) Interpolating schedule s as function of time for a linear schedule (dashed black line) and for the schedule proposed by Roland & Cerf. Time is given in units of the total annealing time τ.

where the items of the database have been promoted to be an orthonormal basis of the Hilbert space $\{1, \ldots, N\} \to \{|1\rangle, \ldots, |N\rangle\}$. Indeed, the marked item $|\overline{m}\rangle$ is the ground state of $\widehat{H}_T$ with energy $E = 0$, while all other items are degenerate excited states with energy $E = J$. Having $\widehat{H}_T$, we generate a QA dynamics by supplementing it with the driving Hamiltonian

$$\widehat{H}_{\rm D} = J\left[1 - |\psi_0\rangle\langle\psi_0|\right], \tag{1.11}$$

where $|\psi_0\rangle = \frac{1}{\sqrt{N}}\sum_{m=1}^{N}|m\rangle$ is both the initial state and the ground state of $\widehat{H}_{\rm D}$. The standard interpolating strategy, given in Eq. (1.1), then leads to

$$\widehat{H}(s) = J\left[s(1 - |\overline{m}\rangle\langle\overline{m}|) - (1-s)(1 - |\psi_0\rangle\langle\psi_0|)\right] \tag{1.12}$$

where, again, the schedule $s(t) \in [0, 1]$ interpolates between $s(0) = 0$ and $s(\tau) = 1$. The simple form of Eq. (1.12), essentially describes a two-level system decoupled from all other states in the Hilbert space. This allows an exact computation of the energy levels of $\widehat{H}(s)$. In Fig 1.3(a), we plot the two lowest energy levels as a function of s. The bottleneck for QA is located at $s = 1/2$ where the gap Δ_s is minimum. Indeed, the analytic expression for the gap Δ_s is

$$\Delta_s^2 = \Delta_{\rm min}^2\left[1 + v^2\left(s - \frac{1}{2}\right)^2\right], \tag{1.13}$$

where $v = 2\sqrt{N-1}$ and $\Delta_{\min} = \frac{J}{\sqrt{N}}$ is the minimum gap encountered in the QA dynamics.

To evaluate the scaling of the running time (or complexity) of the QA algorithm we turn to the adiabatic condition [Albash and Lidar, 2018]

$$\frac{\mathrm{d}t}{\mathrm{d}s} \geq \frac{\hbar}{\epsilon}\frac{||\partial_s H||}{\Delta_s^2} = \frac{\hbar}{J\epsilon} N\sqrt{1-N^{-1}}\frac{\Delta_{\min}^2}{\Delta_s^2} \tag{1.14}$$

where the arbitrary constant ϵ sets the accuracy of the computation. For the simple linear schedule, $\frac{\mathrm{d}t}{\mathrm{d}s} = \tau$ implies that $\tau = \mathcal{O}(N)$. Incidentally, the scaling $\tau \sim \mathcal{O}(N)$ coincides with the result for the complexity of a classical search algorithm. This leads to the disappointing conclusion that a linear-schedule QA gives no improvement over the classical computation!

To recover the expected quantum speed-up, improving over the linear schedule is, therefore, a necessity. In particular, since the adiabatic condition requires a slow passage only in proximity of a vanishing gap, Roland & Cerf suggested to save computational time by choosing a schedule that changes rapidly where the gap is large [Roland and Cerf, 2002] and slows down only close to the critical point. Formally, the improved schedule can be obtained by imposing that the speed $\frac{\mathrm{d}s}{\mathrm{d}t}$ is the maximum allowed by the adiabatic condition given in Eq. (1.14). By replacing the inequality in Eq. (1.14) with an equality, we obtain the differential equation

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{\epsilon}{\hbar}\frac{\Delta_s^2}{||\partial_s H||} = \frac{\epsilon J}{\hbar\sqrt{1-N^{-1}}}\left[1 - 4\frac{N-1}{N}s(1-s)\right] , \tag{1.15}$$

which must be solved with the initial condition $s(0) = 0$. The straightforward integration of this first-order differential equation, imposing that $s(\tau) = 1$, yields

$$s(t) = \frac{1}{2} + \frac{\tan(C(2t/\tau - 1))}{2\tan C} \tag{1.16}$$

where

$$C = \arctan\sqrt{N-1} \tag{1.17}$$

$$\tau = \frac{\hbar}{J\epsilon}\sqrt{N}\arctan\sqrt{N-1}\,. \tag{1.18}$$

This improved schedule $s(t)$ is shown in Fig. 1.3(b). From Eq. (1.18) we read that the complexity of the algorithm is $\tau = \mathcal{O}(\sqrt{N})$, recovering the expected quadratic speed-up.

By improving (or "optimizing") the naif linear schedule, the adiabatic quantum search algorithm gains a quadratic speed-up, which happens to be an essential ingredient in outperforming the classical counterpart of the algorithm. Using this case as an example, we conclude that finding the correct $s(t)$ is an integral part in developing quantum adiabatic algorithms and it should not be neglected.

1.4 Optimal control of the QA schedule

How to devise a practical scheme in which an optimal control of the annealing schedule $s(t)$ is possible? The first step we take in this direction is to get rid of the *continuous* nature of $s(t)$ by switching to a step-discretization of the schedule. Next, we discuss a properly *digitized*-QA, as suggested by Martinis' group [Barends *et al.*, 2016]. This, in turn, leads us to discuss recently proposed hybrid quantum-classical variational schemes on which our solution of the schedule optimization is most naturally framed.

1.4.1 Digitized Quantum Annealing

In many situations it is meaningful, in some cases necessary, to approximate the schedule $s(t)$ by a step function attaining P values $s_1, \cdots, s_{\mathrm{P}}$, with $s_m \in (0, 1]$, corresponding to evolution times Δt_m, with $m = 1, \cdots, \mathrm{P}$ such that $\sum_{m=1}^{\mathrm{P}} \Delta t_m = \tau$. Figure 1.4 is a sketch of such a step-discretization starting from a smooth $s(t)$ — which we might refer to as a *step-QA* scheme — but the discussion below applies to *any* step function. The evolution operator $\widehat{U}(\tau, 0)$ is then given by:

$$\widehat{U}_{\mathrm{QA}}(\tau, 0) \Longrightarrow \widehat{U}_{\mathrm{step}} = \prod_{m=1}^{\leftarrow \mathrm{P}} \mathrm{e}^{-\frac{i}{\hbar}\widehat{H}(s_m)\Delta t_m} \ . \tag{1.19}$$

where the arrow $\leftarrow$ denotes a time-ordered product.

A further digitalization step would be to perform a *Trotter splitting* of the term $\mathrm{e}^{-\frac{i}{\hbar}\widehat{H}(s_m)\Delta t_m}$. To simplify the presentation, here as in the rest of the Thesis[9], we will assume that

$$\widehat{H}_{\mathrm{T}} = \widehat{H}_z \tag{1.20}$$

$$\widehat{H}_{\mathrm{D}} = \widehat{H}_x = -\Gamma \sum_{j=1}^{N} \hat{\sigma}_j^x \ , \tag{1.21}$$

[9] We relax this hypothesis only in Chapter 5, where we consider more general target and driving Hamiltonians.

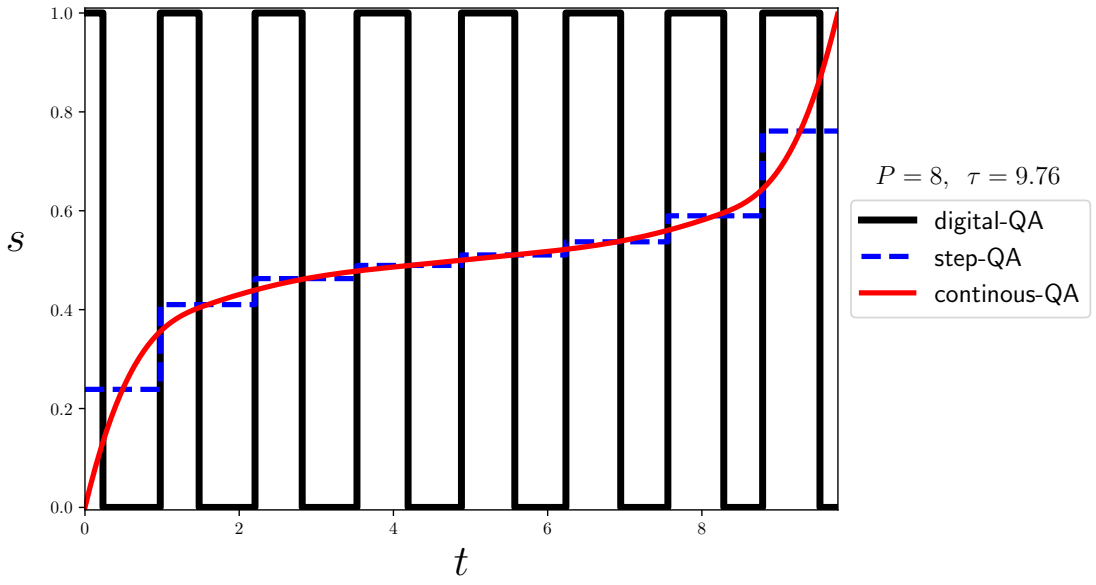

FIGURE 1.4: Digitalization of a smooth schedule $s(t)$. A smooth continuous-time QA $s(t)$, with the associated step-QA and digital-QA schedules. Here $\mathrm{P} = 8$. Notice that the time-intervals Δt_m are not identical. The procedure by which this figure is generated is explained in Ch. 4.

where $\widehat{H}_z$ contains only terms involving $\hat{\sigma}^z$ and $\widehat{H}_x$ contains only terms involving $\hat{\sigma}^x$. Then the lowest-order Trotter splitting:

$$\mathrm{e}^{-\frac{i}{\hbar}\widehat{H}(s_m)\Delta t_m} \simeq \mathrm{e}^{-i\beta_m \widehat{H}_x}\mathrm{e}^{-i\gamma_m \widehat{H}_z} + O((\Delta t_m)^2) \tag{1.22}$$

with

$$\begin{cases} \gamma_m = s_m \dfrac{\Delta t_m}{\hbar} \\ \beta_m = (1 - s_m)\dfrac{\Delta t_m}{\hbar} \end{cases} \tag{1.23}$$

would lead to an approximate evolution operator of the form:

$$\widehat{U}_{\mathrm{QA}}(\tau, 0) \approx \widehat{U}_{\mathrm{digit}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \widehat{U}(\gamma_{\mathrm{P}}, \beta_{\mathrm{P}}) \cdots \widehat{U}(\gamma_1, \beta_1) \ , \tag{1.24}$$

with

$$\widehat{U}(\gamma_m, \beta_m) \equiv \widehat{U}_m = \mathrm{e}^{-i\beta_m \widehat{H}_x}\mathrm{e}^{-i\gamma_m \widehat{H}_z} \ . \tag{1.25}$$

The parameters γ_m and β_m are such that:

$$\sum_{m=1}^{\mathrm{P}} (\gamma_m + \beta_m) = \frac{\tau}{\hbar} \ . \tag{1.26}$$

If the original scheme was a QA based on a smooth $s(t)$, then we might call

this a *digitized-QA* (dQA) [Barends *et al.*, 2016]. A symmetric, or any higher order, Trotter splitting would lead to modified expressions for the parameters $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_{\mathrm{P}})$ and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_{\mathrm{P}})$, with identical sum rule (1.26).

1.4.2 Hybrid Variational schemes and QAOA

Eq. (1.24) naturally leads to the Quantum Approximate Optimization Algorithm (QAOA) algorithm introduced by Farhi *et al.* [2014]. Indeed, one can regard the quantum state

$$|\psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = \widehat{U}(\gamma_{\mathrm{P}}, \beta_{\mathrm{P}}) \cdots \widehat{U}(\gamma_1, \beta_1)|\psi_0\rangle \ , \tag{1.27}$$

attained by the digital evolution operator as variationally dependent on the 2P parameters $(\boldsymbol{\gamma}, \boldsymbol{\beta})$. Using repeated measurements in the computational basis we then evaluate the expectation value of the cost function Hamiltonian

$$E_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle \psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})|\widehat{H}_z|\psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle \ , \tag{1.28}$$

as a function to be minimized through a classical algorithm. The global minimum $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ of the variational parameters determines a corresponding optimal energy $E_{\mathrm{P}}^{\mathrm{opt}} = E_{\mathrm{P}}(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ which is, by construction, a monotonically decreasing function of P. A classical algorithm that uses a quantum processor and a measuring device as an oracle to prepare $|\psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$ and evaluate $E_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})$ is then used to find the values of $\boldsymbol{\gamma}^*$ and $\boldsymbol{\beta}^*$ that minimize E_{P}. Afterward, an approximate solution of the classical problem can be obtained by repeated measures on the state $|\psi_{\mathrm{P}}(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)\rangle$. The total evolution "time" τ, however, is no longer fixed, but rather related to the optimal parameters $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ by the sum rule in Eq. (1.26). For Boolean Satisfiability problems [Garey and Johnson, 1979], where $\widehat{U}_{\mathrm{digit}}(\boldsymbol{\gamma}, \boldsymbol{\beta})$ turns out to be periodic in the variational parameters, one might still show that $\tau < 2\pi\mathrm{P}$. Remarkably Lloyd [2018] showed that the QAOA approach is computationally universal [Lloyd, 2018], although this fact does not guarantee, by itself, efficiency or speed-up [Hastings, 2019].

Incidentally, one can slightly generalize the QAOA construction to look for the ground state of a target Hamiltonian of the form $\widehat{H}_{\mathrm{T}} = \widehat{H}_z + g\widehat{H}_x$. This has been dubbed Variational Quantum-Classical Simulation (VQCS) in Ho and Hsieh [2019].

An interesting recent result [Yang *et al.*, 2017] concerns optimal Quantum Control [D'Alessandro, 2007; Brif *et al.*, 2010]. Indeed, suppose that we fix the total evolution time τ, and one asks for the optimal schedule $s(t)$ with values

bounded in the interval $[0, 1]$ without any continuity or monotonicity requirement. Then, as shown by Yang *et al.* [2017], an application of Pontryagin's principle shows that the optimal schedule has to be of the so-called *bang-bang* form, with $s(t)$ having a square-wave form between the two extremal values 1 and 0, as sketched in Fig. 1.4.

Denoting by γ_m and β_m the "time-lags" spent in the m-th intervals with $s = 1$ and $s = 0$, respectively, we recover once again the form given in Eq. (1.24) or (1.27) where, however, the total number of recursions P is no longer fixed.

1.4.3 Our solution to the problem

In this Thesis we explore a possible solution to the schedule optimization problem. The key idea behind our work is to combine the two well-established quantum optimization schemes we have described, QA and QAOA, to efficiently generate optimal adiabatic ground state preparation protocols without having access to any spectral information (e.g., location and size of the gap).

To achieve this, we go through the path of digitalization, which naturally leads us to consider *digitized*-QA. The schedule optimization problem for *digitized*-QA can then be reformulated in a QAOA framework. In the benchmark case of the Ising chain (and the one dimensional XY model), we show that, combining the QAOA framework with iterative local search algorithms, results in *optimal digitized*-QA protocols that achieve the best performance allowed by Quantum Mechanics. The computational advantage of the *optimal digitized*-QA protocols over QA and QAOA persists when the time needed to find the profile of the schedule is accounted for. This solution is presented in Chaps. 4 and 5.

1.5 Simulated Quantum Annealing

The route we have followed so far exploits a genuine quantum (or hybrid quantum-classical) approach to QA, in the sense that the underlying dynamics is that predicted by the Schrödinger evolution.

However, traditional approaches to QA relay instead on a Markov Chain dynamics to simulate QA on a classical hardware. Indeed, even disregarding the effects of the environment, giving a detailed description of the unitary Schrödinger dynamics of a time-dependent quantum system (see Eq. (1.2)) is a formidable problem. For instance, numerical simulations of the QA dynamics for the SG Hamiltonian are usually limited to very small systems [Kadowaki and Nishimori,

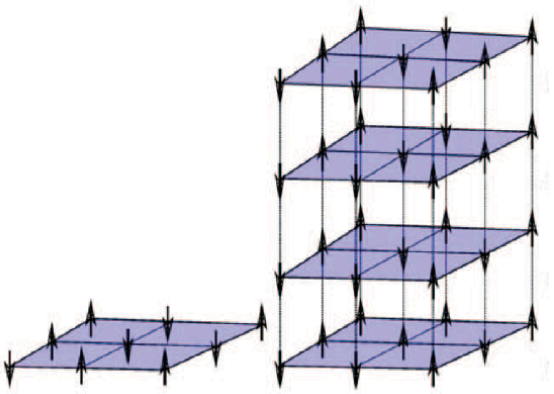

FIGURE 1.5: Illustration of the Path Integral Monte Carlo approach for SQA for a square lattice of $N = 9$ spins, using $\mathrm{P} = 4$ Trotter slices. The left shows the physical system. The right shows the replicas on which SQA is performed. Image from Heim *et al.* [2015].

1998; Farhi *et al.*, 2001], not representative of the actual difficulty of a real problem. This has led, since the early days of QA [Finnila *et al.*, 1994; Santoro *et al.*, 2002], to QA-approaches employing imaginary-time Quantum Monte Carlo (QMC) techniques — most notably Path-Integral Monte Carlo (PIMC) [Ceperley, 1995] and Diffusion Monte Carlo [Becca and Sorella, 2017] —, at least in the most often considered "stoquastic" case, in which off-diagonal matrix elements of the Hamiltonian $\hat{H}(t)$ are non-positive. These approaches are generally known as Simulated QA (SQA), in analogy with classical Simulated Annealing (SA) [Kirkpatrick *et al.*, 1983].

In the case of a PIMC, SQA works as follows [Santoro *et al.*, 2002; Martoňák *et al.*, 2002]: one simulates the quantum system in Eq. (1.1) at a fixed s and temperature T, corresponding to $\beta = 1/(k_B T)$ by resorting to a Suzuki-Trotter path-integral [Suzuki, 1976], which involves mapping the equilibrium quantum partition function

$$Z_Q = \mathrm{Tr}\left[\mathrm{e}^{-\beta \widehat{H}(s)}\right] , \tag{1.29}$$

into the partition function of an equivalent classical Ising system with P coupled replicas of the original lattice, as shown in Fig. 1.5. In principle one should take $\mathrm{P} \to \infty$, a limit in which the mapping becomes exact. Then, during the SQA simulation, the value of $s(t)$ is increased step-wise as a function of the *Monte Carlo time* t going from $s(0) = 0$ to $s(\tau) = 1$.

Recent results on SQA have consolidated it as a powerful quantum-inspired optimization algorithm [Crosson and Harrow, 2016; Baldassi and Zecchina, 2018].

Indeed, quite remarkably, as shown in Baldassi and Zecchina [2018], non-convex optimization problems are known in which SQA, with the $P \to \infty$ limit properly taken, is more efficient than its classical SA counterpart.

However, the SQA approach still raises two issues. On the one hand, SQA is built on a classical Markov-chain dynamics which is in principle *unrelated* to the Schrödinger quantum dynamics of a real QA device. On the other hand, the Suzuki-Trotter imaginary-time discretization would require taking the so-called *time-continuum limit* $P \to \infty$ [Heim *et al.*, 2015]; however, regarding SQA as a classical optimization algorithm, one might be interested in finding the optimal value of P [Santoro *et al.*, 2002].

Concerning the Monte Carlo vs. physical dynamics issue, some evidence on ground state success probability histograms for Ising problems [Boixo *et al.*, 2014] encouraged to believe that SQA might have something to do with the actual physical dynamics of a real-world QA hardware: indeed, a certain degree of correlation between the performance of SQA and that of the D-Wave One QA device on random Ising instances with $N = 108$ qubits were found. Equally encouraging was the message of Isakov *et al.* [2016] (see also Mazzola *et al.* [2017]) on the tunneling rate between the two ground states of an ordered Ising ferromagnet: indeed, a correlation between the size-scaling of the PIMC tunneling rate and the incoherent tunneling rate of a physical device was found. Finally, Denchev *et al.* [2016] demonstrated a possible role of finite-range tunneling on specially crafted random Ising instances; for those, the results obtained by SQA appeared to have a scaling with problem size which is similar to that obtained by the D-Wave 2X annealer, albeit with a large constant overhead; on the contrary, the size scaling of SA was poorer.

Concerning the time-continuum limit issue, Heim *et al.* [2015] have pointed out that the optimization advantage of PIMC-SQA against classical SA, observed in Santoro *et al.* [2002] for a suitably optimal finite value of P in a two-dimensional random Ising model, might disappear when the limit $P \to \infty$ is properly taken.

The two issues are in some way related, since, quite evidently, the role of the $P \to \infty$ limit loses part of its meaning unless the SQA dynamics has something to do with the actual *physical* dynamics.

1.6 Thesis Outline

The present Thesis dwells with some of the issues that we mentioned in the previous introductory presentation.

We start, in Chap. 2, from the issues arising when attempting to simulate QA on a classical computer. Here, we re-examine the relationship between the SQA Monte-Carlo dynamics and the actual Schrödinger dynamics. We do that in the specific case of the disordered quantum Ising chain, where analytical tools allow us to simulate the Schrödinger dynamics of large systems. The results presented in this chapter are published in [Mbeng *et al.*, 2019]

Next, we move to a series of studies that provide a link between traditional QA and alternative hybrid quantum-classical approaches, with emphasis on optimal quantum control and schedule optimization. Our contribution to the field is twofold.

As a first contribution, we develop a technique to establish a variational bound on the performance of QAOA for translationally invariant spin problems on regular periodic graphs. We do this by exploiting the intrinsic flexibility in the boundary conditions of a reduced spin problem to prove rigorous bounds on the relative error $\epsilon_{\mathrm{P}}^{\mathrm{res}}$ of the QAOA optimal solution on a circuit of depth P. We first present the technique in one dimension, see Chap. 4, and then extended to a more general setting and a continuous-time dynamics, see Chap. 5. When applied to the translationally invariant quantum Ising chain, which is analytically tractable, we show that a quadratic speed-up with respect to the standard linear-QA annealing is allowed.

As a second contribution, we establish a link between optimal quantum control and the adiabatic dynamics behind *digitized*-QA [Barends *et al.*, 2016]. Indeed, using the QAOA framework, we explicitly show that one can iteratively single-out a *smooth regular* solution which can be regarded as an *optimal digitized*-QA schedule, satisfying all the expected requirements for adiabaticity in a digital context, without any need for spectral information on the Hamiltonian. Moreover, the construction of the *optimal digitized*-QA schedule turns out to be computationally less expensive than searching for unstructured QAOA solutions.

Here is a layout of the Thesis. We start in Chap. 2 by comparing the SQA Monte-Carlo dynamics and the QA Schrödinger dynamics on a random Ising chain. Next, in Chap. 3 we give an analysis of the optimal working point for linear schedules in *digitized*-QA. Using analytical and numerical tools, we show that there is an optimal annealing time $\tau_{\mathrm{P}}^{\mathrm{opt}}$ beyond which *digitized*-QA performs poorly. In

Chap. 4 we show how a *digitized*-QA can be made optimal without any prior knowledge on the location of the spectral gap, and we introduce the technique to establish variational bounds on its performance. Chap. 5 contains the generalization of these ideas to higher-dimensional problems and more general ground-state preparation problems. Chap. 6, finally, contains a summary of the results and a conclusive discussion. The Thesis ends with five Appendices that contain more technical material.

Chapter 2

Simulated Quantum Annealing: Monte Carlo vs Schrödinger dynamics

Simulated Quantum Annealing (SQA) is a classical computational strategy that emulates a Quantum Annealing (QA) dynamics through a Path Integral Monte Carlo (PIMC), whose parameters are changed during the simulation. In this Chapter we apply SQA to the transverse field Ising chain. The explicit expression of the target Hamiltonian is

$$\widehat{H}_{\mathrm{T}} = \widehat{H}_z = -\sum_{i=1}^{N-1} J_i \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \ , \tag{2.1}$$

where $\hat{\sigma}_i^{x,y,z}$ are Pauli matrices at site i, $J_i > 0$ are positive couplings and N is the number of spins in the chain. Following the QA approach, we supplement $\widehat{H}_z$ with the standard transverse field driving:

$$\widehat{H}_{\mathrm{D}} = \widehat{H}_x = -\sum_{j=1}^{N} \hat{\sigma}_j^x \ . \tag{2.2}$$

Under these circumstances, a possible measure [Santoro *et al.*, 2002] of the accuracy of the ground state preparation protocol is the density of defects

$$\rho_{\mathrm{def}}(\tau) = \frac{1}{2N} \sum_{i=1}^{N-1} \langle \psi(\tau) | 1 - \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z | \psi(\tau) \rangle \ , \tag{2.3}$$

where τ is the annealing time and $|\psi(\tau)\rangle$ is the output state of the annealing

protocol. Indeed, the density of defects $\rho_{\rm def}(\tau)$ is proportional to the average Hamming[1] distance between the target state and the output state of the algorithm.

Previous works [Caneva *et al.*, 2007; Zanca and Santoro, 2016] have shown that, in the presence of disorder, a coherent QA provides a quadratic speedup with respect to classical Simulated Annealing (SA), with a density of Kibble-Zurek (KZ) defects decaying as $\rho_{\rm KZ}^{\rm QA} \sim (\log \tau)^{-2}$ as opposed to $\rho_{\rm KZ}^{\rm SA} \sim (\log \tau)^{-1}$ — τ being the total annealing time — while for the translationally invariant case both give the same power-law $\rho_{\rm KZ}^{\rm QA} \approx \rho_{\rm KZ}^{\rm SA} \sim \tau^{-1/2}$. We show that the dynamics of SQA, while correctly capturing the Kibble-Zurek scaling $\tau^{-1/2}$ for the translationally invariant case, is unable to reproduce the QA dynamics in the disordered case at intermediate τ. We analyze and discuss several issues related to the choice of the Monte Carlo moves (local or global in space), the time-continuum limit needed to eliminate the Trotter-discretization error, and the long auto-correlation times shown by a local-in-space Monte Carlo dynamics for large disordered samples.

2.1 Introduction

The SQA approach raises two important issues. This first issue is that the classical Markov-chain (Monte Carlo) dynamics implemented by SQA is in principle *unrelated* to the Schrödinger quantum dynamics of a real QA device. The second issue is related to the Suzuki-Trotter imaginary-time discretization which would require taking the *time-continuum limit* $\mathrm{P} \to \infty$. We recall that the two issues are deeply connected, as the $\mathrm{P} \to \infty$ limit is relevant only if the SQA dynamics can mimic the behavior of the actual QA *coherent* dynamics.

Here, we will reconsider these issues, trying to shed light on some aspects of the Monte Carlo dynamics behind PIMC-SQA. We decided to concentrate our efforts on a detailed analysis of the *random Ising chain*, where exact benchmarks, both for thermal equilibrium and coherent-QA results, are readily obtained by a Jordan-Wigner [Lieb *et al.*, 1961; Young and Rieger, 1996] mapping to free fermions, and SA results are also available [Suzuki, 2009; Zanca and Santoro, 2016]. The coherent QA dynamics of a transverse-field random Ising chain was extensively studied in Dziarmaga [2006], Caneva *et al.* [2007], and Zanca and Santoro [2016]. It is known that the Kibble-Zurek density of defects generated by adiabatically crossing [Kibble, 1980; Zurek, 1996; Polkovnikov *et al.*, 2011] the

[1] The Hamming distance between two strings of equal length is the number of positions at which the corresponding symbols are different.

$T = 0$ quantum critical point [Fisher, 1995] scales as $\rho_{\rm KZ}^{\rm QA} \sim (\log \tau)^{-2}$, hence with a quadratic speedup with respect to the Huse-Fisher bound [Huse and Fisher, 1986] for SA on the same model, $\rho_{\rm KZ}^{\rm SA} \sim (\log \tau)^{-1}$; the SA bound is numerically confirmed by Monte Carlo SA [Suzuki, 2009], and by deterministic single spin-flip classical master equation SA [Zanca and Santoro, 2016]. Hence, a *limited quantum speed-up*, in the terminology of Rønnow *et al.* [2014], is well documented for the random Ising chain, although the classical ground states are simply the two trivial ferromagnetic states. In the absence of disorder, on the contrary, there is no limited quantum speedup: the Kibble-Zurek mechanism leads, for both QA [Dziarmaga, 2005] and SA [Suzuki, 2009; Zanca and Santoro, 2016], to a power-law scaling of the density of defects: $\rho_{\rm KZ}^{\rm QA} \approx \rho_{\rm KZ}^{\rm SA} \sim \tau^{-1/2}$.

Our results can be summarized as follows. Due to the absence of frustration, we can compare PIMC-SQA results obtained with two types of Monte Carlo moves: Swendsen-Wang [Swendsen and Wang, 1987] (SW) cluster moves limited to the imaginary-time direction, hence local in space, with space-time (non-local) SW cluster moves, which provides an extremely fast Monte Carlo dynamics. We find that equilibrium thermodynamical PIMC simulations at finite T clearly show a sampling problem, more precisely, large Monte Carlo auto-correlation times, emerging for large P when local SW cluster moves limited to the time-direction — the most natural candidate moves for a "physical" single-spin-flip dynamics — are employed below the critical point $\Gamma < \Gamma_c$, at low temperatures and for large sample sizes N. Next, we compare the annealing dynamics of SQA with coherent-QA evolution results performed by solving the time-dependent Bogoljubov-de Gennes equations for the Jordan-Wigner fermions [Caneva *et al.*, 2007]. We show that, while the standard Kibble-Zurek $\tau^{-1/2}$ scaling [Kibble, 1980; Zurek, 1996; Polkovnikov *et al.*, 2011] of the residual energy is recovered in the ordered case, in the presence of disorder the situation is more complicated. The SQA dynamics shows a very interesting feature: the residual energy at $\Gamma(t)$ is essentially predicted by the corresponding *equilibrium* thermodynamical value, but at an effective temperature $T_{\rm eff}(\tau) > T$. This aspect is shared by the coherent-QA evolutions, which can also be described by a similar *Ansatz*. However, the overall behaviors of $T_{\rm eff}(\tau)$ in the two cases, or equivalently that of $\varepsilon_{\rm res}(\tau)$ vs. τ, are unrelated. In particular, at intermediate times τ, the SQA data show a smaller power-law slope as compared to the coherent-QA data.

The rest of the Chapter is organized as follows. Section 2.2 presents the model we study, the random Ising chain in a transverse field, and briefly describes the

methods used: exact Jordan-Wigner mapping to free fermions and PIMC. Section 2.3 contains our results, both at equilibrium (Sec. 2.3.1) and for QA (Sec. 2.3.2). Section 2.4, finally, contains our concluding remarks.

2.2 Model and methods

We consider a random Ising model in one dimension (1D) with open boundary conditions with $\widehat{H}_z$ and $\widehat{H}_{\rm D}$ given in Eq. (2.1) and Eq. (2.2). The QA interpolating Hamiltonian is[2]

$$\widehat{H}(\Gamma) = -\sum_{i=1}^{N-1} J_i \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z - \Gamma \sum_{i=1}^{N} \hat{\sigma}_i^x \ . \tag{2.4}$$

We assume the bond couplings J_i to be uniformly distributed independent positive random variables, $J_i \in (0, 1]$. For $\Gamma = 0$, because of the one-dimensional geometry of the system, disorder causes no frustration, and the optimization task of finding the classical ground state of $\widehat{H}_z$ is trivial: the two degenerate classical ground states of the system are simply the ferromagnetic states $|\uparrow\uparrow\cdots\uparrow\rangle$ and $|\downarrow\downarrow\cdots\downarrow\rangle$, with a minimum energy (per spin), given by $\epsilon_{\rm gs}(\Gamma = 0) = -\frac{1}{N}\sum_{i=1}^{N-1} J_i$. Nevertheless, disorder alone is sufficient to make the annealing *dynamics* — both classical [Suzuki, 2009; Zanca and Santoro, 2016] and quantum [Dziarmaga, 2006; Caneva *et al.*, 2007; Zanca and Santoro, 2016] — rather complex.

Once again, PIMC is a standard approach to simulate the *equilibrium* properties of the Hamiltonian (2.4) at finite temperature $T > 0$ when Γ does not depend on time. It works as follows. We first apply a standard Suzuki-Trotter [Suzuki, 1976] mapping of the quantum system at a fixed temperature T, corresponding to $\beta = 1/(k_B T)$, into $\mathrm{P} \to \infty$ classical coupled replicas:

$$Z_Q = \mathrm{Tr}\, \mathrm{e}^{-\beta \widehat{H}} \simeq \lim_{\mathrm{P}\to\infty} \sum_{S}^{\text{config}} \mathrm{e}^{-K_{\rm cl}[S]} \ , \tag{2.5}$$

which interact with a classical action

$$K_{\rm cl} = -\beta_{\rm P} \sum_{k=1}^{\mathrm{P}} \sum_{i=1}^{N-1} \left(J_i \, S_i^k \, S_{i+1}^k + J^{\perp} S_i^k \, S_i^{k+1} \right) \ , \tag{2.6}$$

at an effective temperature $\mathrm{P}T$, corresponding to $\beta_{\rm P} \equiv \beta/\mathrm{P} \equiv \Delta\tau$. Here $S_i^k = \pm 1$ is a classical Ising spin at site i and "imaginary-time slice" $\tau_k = (k-1)\Delta\tau$, with

[2] In this chapter we find it convenient to use a slightly different notation for the QA Hamiltonian than the one given in Eq. (1.1).

$k = 1 \cdots \mathrm{P}$ and boundary condition $S_i^{\mathrm{P}+1} \equiv S_i^1$, as required by the quantum trace in the partition function. (The sum over configurations in Eq. (2.5) runs over $S = \{S_i^k\}$.) The uniform ferromagnetic coupling $J^\perp$ along the imaginary-time direction is set by:

$$J^\perp = -\frac{1}{2\beta_{\mathrm{P}}} \log\left[\tanh\left(\beta_{\mathrm{P}}\Gamma\right)\right] \ . \tag{2.7}$$

The correct quantum mechanical equilibrium calculation is recovered by taking the limit $\mathrm{P} \to \infty$. Using a Metropolis algorithm we can then implement several different Monte Carlo dynamics for K_{cl}, depending on the choice of the *Monte Carlo moves* on which the corresponding classical Markov chain is built. In an equilibrium PIMC, this would make no difference for the final equilibrium averages: it would just influence how fast the system reaches the equilibrium steady state on which averages are calculated. In an annealing framework, the choice of the Monte Carlo moves is a delicate matter influencing the outcome of the SQA simulation. Indeed, SQA is built by appropriately changing the transverse field Γ during the course of the PIMC simulation in the hope of mimicking the physical annealing dynamics behind Eq. (2.4): there is no intrinsic separation between transient and stationary state. In the following, we will investigate and compare two different Monte Carlo moves:

1) *time cluster flips* (local in space). Given a site i, clusters of spins $\{S_i^k\}$ are constructed along the imaginary-time direction using the Swendsen-Wang algorithm [Swendsen and Wang, 1987]. This is the choice of Heim *et al.* [2015]. A single Monte Carlo Sweep (MCS) consists of N time-cluster flips.

2) *space-time cluster flips* (non-local). Since the classical action in Eq. (2.6) is ferromagnetic (unfrustrated), one can adopt algorithms which construct space-time clusters, either Swendsen-Wang [Swendsen and Wang, 1987] or Wolff [Wolff, 1989]. A MCS consists of a single space-time cluster flip.

The advantage of working with a random Ising chain is that *exact* equilibrium as well as coherent evolution QA results can be easily obtained and compared to PIMC data. Indeed, using a Jordan-Wigner transformation, the Hamiltonian in Eq. (2.4) can be mapped to the following free-fermion Hamiltonian (see App. A)

$$\widehat{H} = -\sum_{i=1}^{N-1} J_i(\hat{c}_{i+1}^\dagger \hat{c}_i + \hat{c}_i^\dagger \hat{c}_{i+1}^\dagger + \mathrm{H.c.}) - \Gamma \sum_{i=1}^{N} (2\hat{c}_i^\dagger \hat{c}_i - 1) \ , \tag{2.8}$$

where $\hat{c}_i^\dagger$ and $\hat{c}_i$ are spinless fermionic operators. In equilibrium — Γ independent of t — one can diagonalize such a BCS-like Hamiltonian by a Bogoliubov transformation, constructed by the numerical diagonalization of a $2N \times 2N$ matrix [Young and Rieger, 1996; Caneva *et al.*, 2007]. The relevant quantity that we will consider is the difference (per spin) between the interaction energy's thermal average at a given value of Γ and T, and the ferromagnetic classical ground-state energy $\epsilon_{\rm gs}(\Gamma = 0)$:

$$\varepsilon_{\rm c}(\Gamma, T) = \frac{1}{N} \sum_{i=1}^{N-1} J_i \left(1 - \left\langle \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \right\rangle_{\Gamma,T}\right) . \tag{2.9}$$

$\varepsilon_{\rm c}(\Gamma, T)$ quantifies thermal and quantum fluctuations over the classical ground states energy. Within a coherent-QA framework, where $\Gamma(t)$ is slowly switched to 0 in a timescale τ and the Schrödinger dynamics (1.2) is followed, one can evaluate the time-dependent residual energy:

$$\epsilon_{\rm res}(t, \tau) = \frac{1}{N} \sum_{i=1}^{N-1} J_i \left(1 - \left\langle \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \right\rangle_t\right) , \tag{2.10}$$

where now $\left\langle \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \right\rangle_t = \langle \psi(t) | \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z | \psi(t) \rangle$ is the quantum average with the time-evolving state $|\psi(t)\rangle$. It can be calculated through time-dependent Bogoljoubov-de Gennes (BdG) equations [Caneva *et al.*, 2007; Zanca and Santoro, 2016]. The residual energy at the end of the annealing is simply obtained as

$$\varepsilon_{\rm res}(\tau) \equiv \epsilon_{\rm res}(t = \tau, \tau) . \tag{2.11}$$

2.3 Results

We now discuss the results obtained on the random Ising chain problem. We start from the equilibrium thermodynamics at finite T and Γ, where we compare how different choices of Monte Carlo moves — specifically, Swendsen-Wang cluster moves restricted to the time direction only, or extended to space and time — affect the convergence to exact results for $\mathrm{P} \to \infty$. Interestingly, we find that, in the presence of disorder, there is a clear *sampling problem*, associated to large Monte Carlo auto-correlation tails, for the time cluster moves as P increases in the ferromagnetic phase $\Gamma < \Gamma_c$, for large enough N.

Next, we compare the annealing dynamics of SQA with coherent-QA evolution results performed by solving the time-dependent Bogoljubov-de Gennes equations for the Jordan-Wigner fermions [Caneva *et al.*, 2007]. We show that,

while SQA recovers the standard Kibble-Zurek $\tau^{-1/2}$ scaling [Kibble, 1980; Zurek, 1996; Polkovnikov *et al.*, 2011] of the residual energy in the ordered case, in the presence of disorder, the situation is less clear.

2.3.1 Equilibrium PIMC simulations

Figure 2.1 shows our PIMC equilibrium estimates for $\varepsilon_{\rm c}(\Gamma, T)$ in Eq. (2.9) at low temperature, $T = 0.01$, for two values of Γ, above and below the $T = 0$ quantum critical point [Fisher, 1995], here at $\Gamma_c = 1/\mathrm{e}$ — where e is the Napier's number — with $J_i \in [0, 1]$, which sets our energy scale. The data shown refer to a *single realization* of disorder (no disorder average), in order to precisely test their convergence to the exact value for the same realization of $\{J_i\}$: they are representative of all the instances we have tested. Results for both types of Monte Carlo moves are shown by triangles (time cluster moves) and squares (space-time cluster moves). For $\Gamma = 1 > \Gamma_c$, we see that both Monte Carlo moves provide consistent estimates of $\varepsilon_{\rm c}(\Gamma, T)$, which approach from below the correct exact value, denoted by the horizontal line, as $\mathrm{P} \to \infty$. Notice that, for finite P, the Trotter discretization error — of order $\mathcal{O}(\frac{1}{\mathrm{P}^2 T^3})$, which amounts to a 10% error at $\mathrm{P} \approx 100$ for the case shown —, introduces a bias towards lower values of $\varepsilon_{\rm c}$. Even more interesting is the outcome for $\Gamma = 0.1 < \Gamma_c$, see Fig. 2.1 (b). Here we see that the space-time (non-local) cluster moves correctly reproduce the exact $\mathrm{P} \to \infty$ value, with the usual Trotter-error bias. However, the time-cluster moves (local in space) *completely miss the exact target*: as P increases, the PIMC value first seems to move towards the exact one, up to about $\mathrm{P}_* \sim 32 \div 64$, but then strongly overshoots the target and shows deviations as large as a 100% error for the highest $\mathrm{P} = 1024$. To elucidate more this sampling difficulty, we have evaluated the connected auto-correlation function for the Monte Carlo estimate of ε_c:

$$C(\Delta t) = \frac{\langle \varepsilon_c(t + \Delta t)\varepsilon_c(t)\rangle - \langle \varepsilon_c(t + \Delta t)\rangle \langle \varepsilon_c(t)\rangle}{\langle \varepsilon_c^2(t)\rangle - \langle \varepsilon_c(t)\rangle^2} , \tag{2.12}$$

where t is a Monte Carlo time and $\langle \cdots \rangle$ denotes an average over t. Figure 2.2 shows the results obtained for $C(\Delta t)$ in the regime where the sampling is difficult, i.e., for $\Gamma < \Gamma_c$, at small T and large $\mathrm{P} = 512$. The data refer to different sample sizes N obtained by cutting a central portion from a single realization of J_i at $N = 128$. The decay of the auto-correlations is exponential for small sizes $N < 64$, as highlighted in the inset of Fig. 2.2, but develops large tails for larger N, see data for $N = 128$. This implies that the time cluster moves are

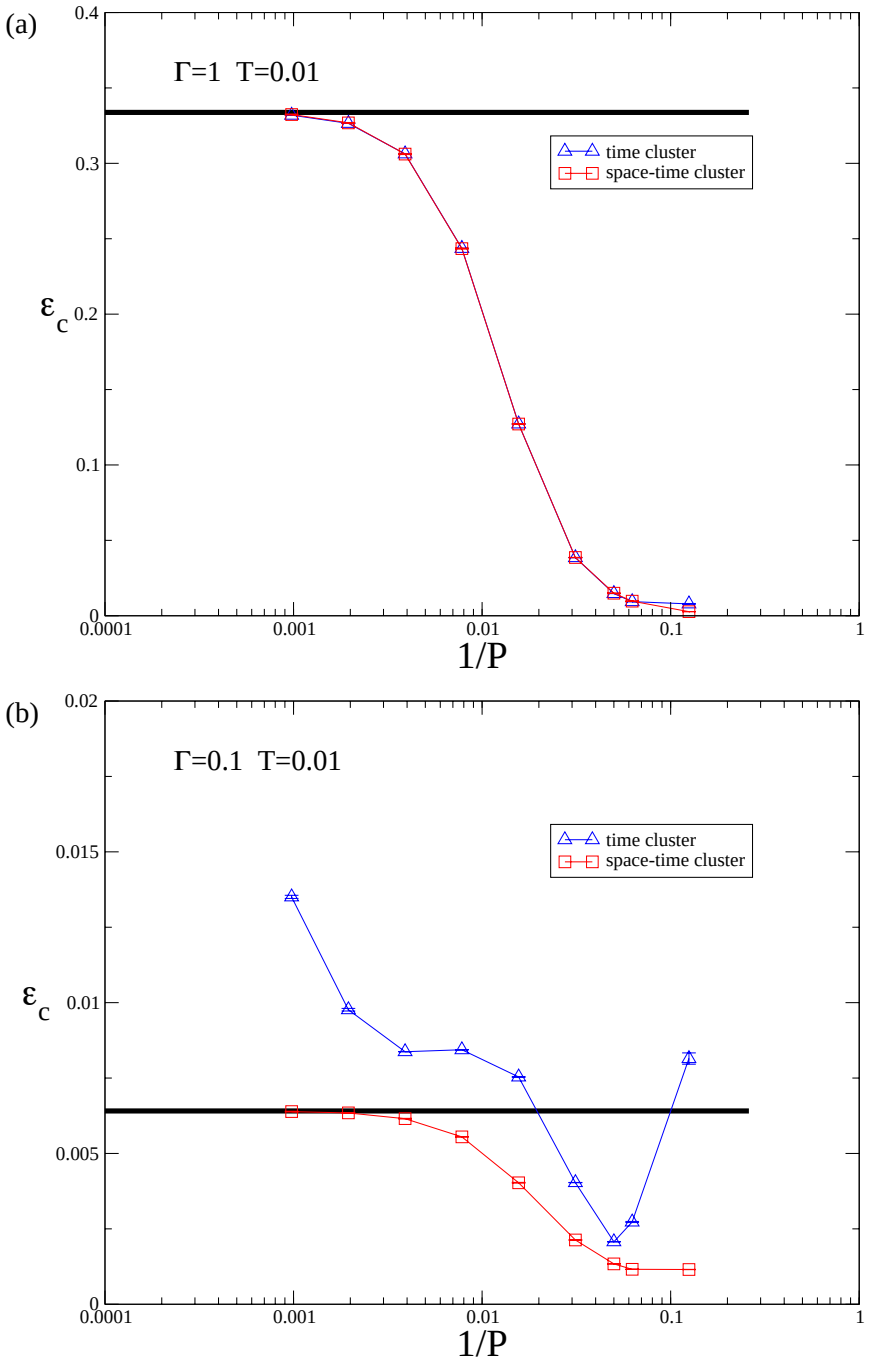

FIGURE 2.1: PIMC estimates of $\varepsilon_c(\Gamma, T)$ in Eq. (2.9) at low temperature $k_B T = 0.01$ and for two values of the transverse field, $\Gamma = 1$ (a) and $\Gamma = 0.1$ (b), as a function of the inverse number of Trotter slices 1/P. The results are obtained for a given random instance of a chain of $N = 256$ spins. The thick horizontal lines denote the exact $\varepsilon_c(\Gamma, T)$ calculated using the Jordan-Wigner calculation. The simulation length is $t_{\rm run} \lesssim 10^8$MCS. An initial $t_{\rm burn}$ MCS were discarded to ensure the equilibration of the Markov chain. The value of $t_{\rm burn}$ was chosen using Geweke's diagnostic [Brooks and Roberts, 1998], while taking care to not discard more than 50% of the iterations.

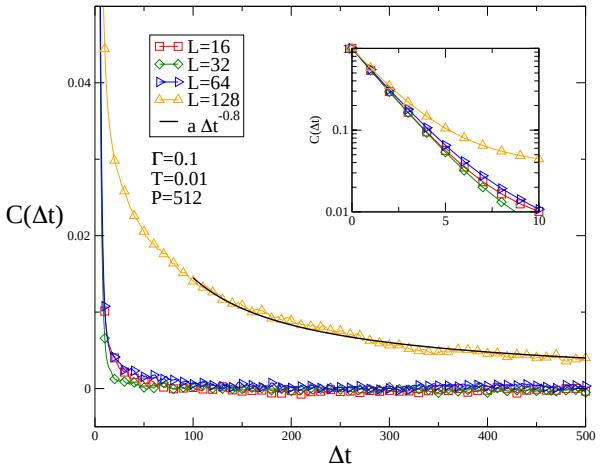

FIGURE 2.2: The connected time-correlation function $C(\Delta t)$ for $\varepsilon_c(\Gamma, T)$, as defined in Eq. (2.12). The averages are calculated with PIMC using the time-cluster moves. Each curve is computed using a single PIMC simulation of length $t_{\text{run}} = 10^8$MCS. To ensure equilibration, we discard the first half of each Markov chain ($t_{\text{burn}} = 5 \cdot 10^7$MCS). The data for $N = 128$ are compatible with a power-law decay of auto-correlations. The inset highlight the initial exponential decay of correlations for small N.

unable to correctly sample the correct distribution, especially at relatively large P and N, even with quite long simulation times of order $t_{\text{run}} \sim 10^8$ MCS. The fact that a large-P Trotter sampling is definitely non-trivial is well known for PIMC in continuous systems, see for instance Brualla *et al.* [2004]. The time cluster moves are, however, the only candidate moves for PIMC in frustrated systems, where space-time cluster moves cannot be employed; they are also a quite natural implementation of a "physical" single-spin-flip dynamics.

It should be noted that the mere fact that the time-cluster PIMC shows a kind of "dynamical slowing down" does not, by itself, imply that a PIMC-SQA based on it would not be able to capture the physical Kibble-Zurek QA behavior. Indeed, on the contrary, we will show that the space-time cluster moves, undoubtedly very efficient in sampling the equilibrium behavior in all regimes, are *dynamically unphysical*: they are unable to capture, for instance, the Kibble-Zurek scaling of the ordered Ising chain density of defects.

2.3.2 PIMC-SQA compared to coherent QA

We now turn to the SQA dynamics. As done in previous studies [Santoro *et al.*, 2002; Martoňák *et al.*, 2002; Heim *et al.*, 2015], we use a protocol in which Γ is linearly reduced to zero as a function of the Monte Carlo time. More precisely, we start from $\Gamma(0) = 2.5$ and perform a preliminary equilibration of the system. We

then reduce $\Gamma(t)$ at each MCS in such a way that $\Gamma(\tau) = 0$, where t is the time in MCS units and τ the total annealing time:

$$\Gamma(t) = \Gamma(0) \left(1 - \frac{t}{\tau}\right) . \tag{2.13}$$

Notice that, in our choice, we reduce Γ at each MCS, by a rather small quantity $\Delta\Gamma = \Gamma(0)/\tau$, rather than implementing a staircase with N_Γ MCS at each of the τ/N_Γ values of Γ: the results are essentially equivalent. Let us consider, as a warm up, the ordered case $J_i = J$, where we set $J = 1$ to be our energy unit. The coherent QA dynamics [Dziarmaga, 2005] here is well known to produce a Kibble-Zurek [Kibble, 1980; Zurek, 1996; Polkovnikov *et al.*, 2011] decrease of the residual energy $\varepsilon_{\rm res}(\tau) \sim \tau^{-1/2}$. The SQA estimate, calculated from the Trotter-replica average

$$\varepsilon_{\rm res}^{\rm avg}(\tau) = \frac{1}{N} \sum_{i=1}^{N-1} J_i \left(1 - \frac{1}{P} \sum_{k=1}^{P} S_i^k S_{i+1}^k\right) , \tag{2.14}$$

at the final configuration $\{S_i^k\}$, averaged over many repetitions of the SQA run, is shown in Fig. 2.3 for the time cluster moves. Quite remarkably, the behavior of $\varepsilon_{\rm res}^{\rm avg}(\tau)$ is compatible with the KZ coherent behavior, with the large τ deviations likely due to finite-size effects. As already anticipated, we have verified that such KZ scaling would not be captured by using an SQA dynamics based on non-local (space-time) cluster moves.

The natural question is whether this agreement also survives in the disordered case. We start by showing the results obtained, in the same spirit of the SQA numerics presented in Santoro *et al.* [2002]; Martoňák *et al.* [2002]; Heim *et al.* [2015], by considering the Trotter slice $k_\star$ that realizes the minimum classical energy value for the residual energy:

$$\varepsilon_{\rm res}^{\rm opt}(\tau) = \frac{1}{N} \sum_{i=1}^{N-1} J_i \left(1 - S_i^{k_\star} S_{i+1}^{k_\star}\right) , \tag{2.15}$$

for a given random instance of a chain with $N = 256$ sites and $J_i \in [0, 1]$. In Fig. 2.4(a) we show SQA data obtained for various P with the SW time cluster moves. Notice the strong similarity with the SQA data shown in Fig. 2.4(b) from Santoro *et al.* [2002] and, in particular, with Fig. 2.4(c) from Heim *et al.* [2015], obtained for a two-dimensional frustrated Ising glass: this shows that quite likely, the phenomena observed are due to disorder, rather than to a truly complex frustrated landscape. Notice also that, within an optimization framework, the

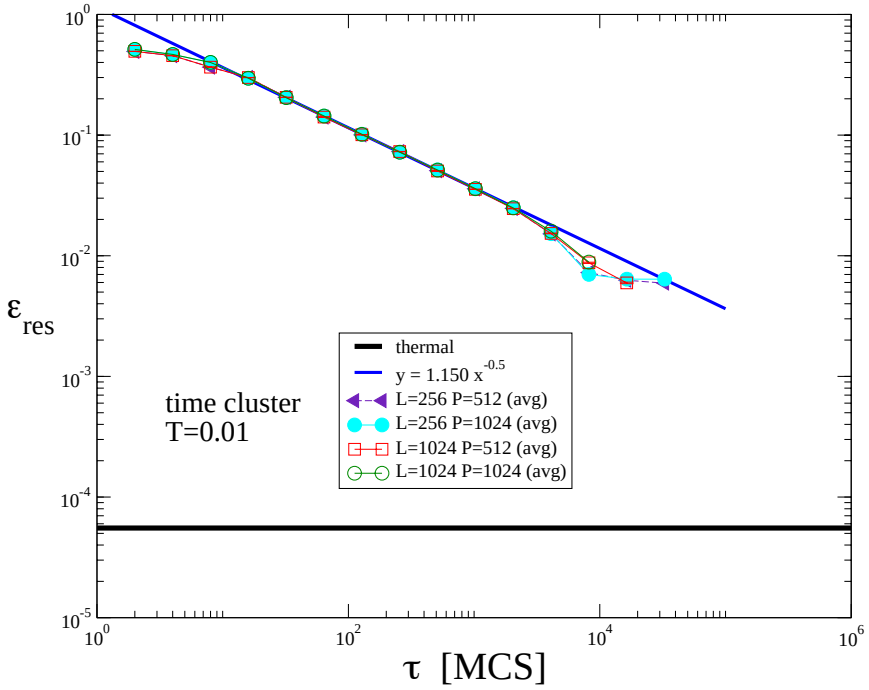

FIGURE 2.3: Test for Kibble-Zurek behaviour, $\varepsilon_{\rm res}(\tau) \sim \tau^{-1/2}$, in the ordered transverse-field Ising model. SQA is here implemented with SW time cluster moves at $T = 0.01$. The horizontal thick line denotes the equilibrium thermal value of $\varepsilon_{\rm c}(\Gamma = 0, T = 0.01)$.

optimal choice of P is not $\mathrm{P} \to \infty$, but rather $\mathrm{P}_{\rm opt} \sim 32$, as indeed empirically found in Santoro *et al.* [2002]. As pointed out in Heim *et al.* [2015], these results raise doubts as to whether any possible advantage of SQA over plain SA might be lost in the proper quantum limit $\mathrm{P} \to \infty$. Figure 2.5 shows that the SQA results obtained with the SW space-time cluster moves behave in a completely different way: they quickly converge to the expected thermal average $\varepsilon_{\rm c}(\Gamma = 0, T = 0.01)$. This tells us that the SQA results are highly sensitive to the type of MC moves one adopts, as perhaps expected: most likely, the space-time cluster non-local moves have little to do with any physical dynamics, as we will further comment on in the following.

Returning to the time cluster SQA results, we re-plot them in the inset of Fig. 2.4(a) for the largest P, to highlight the fact that the $\mathrm{P} \to \infty$ limit is indeed reached as soon as $\mathrm{P} \geq 256$. Here, the two sets of data shown are $\varepsilon_{\rm res}^{\rm opt}(\tau)$, the optimal Trotter-slice value in Eq. (2.15), and the proper "quantum average" $\varepsilon_{\rm res}^{\rm avg}(\tau)$ in Eq. (2.14), which shows a much smoother and monotonic behaviour: notice that the two curves approach each other for the largest τ investigated. This witnesses the fact that, for these largest τ, the quantum fluctuations — *i.e.* the fluctuations along the Trotter-time direction — seem to play no role towards the end of the

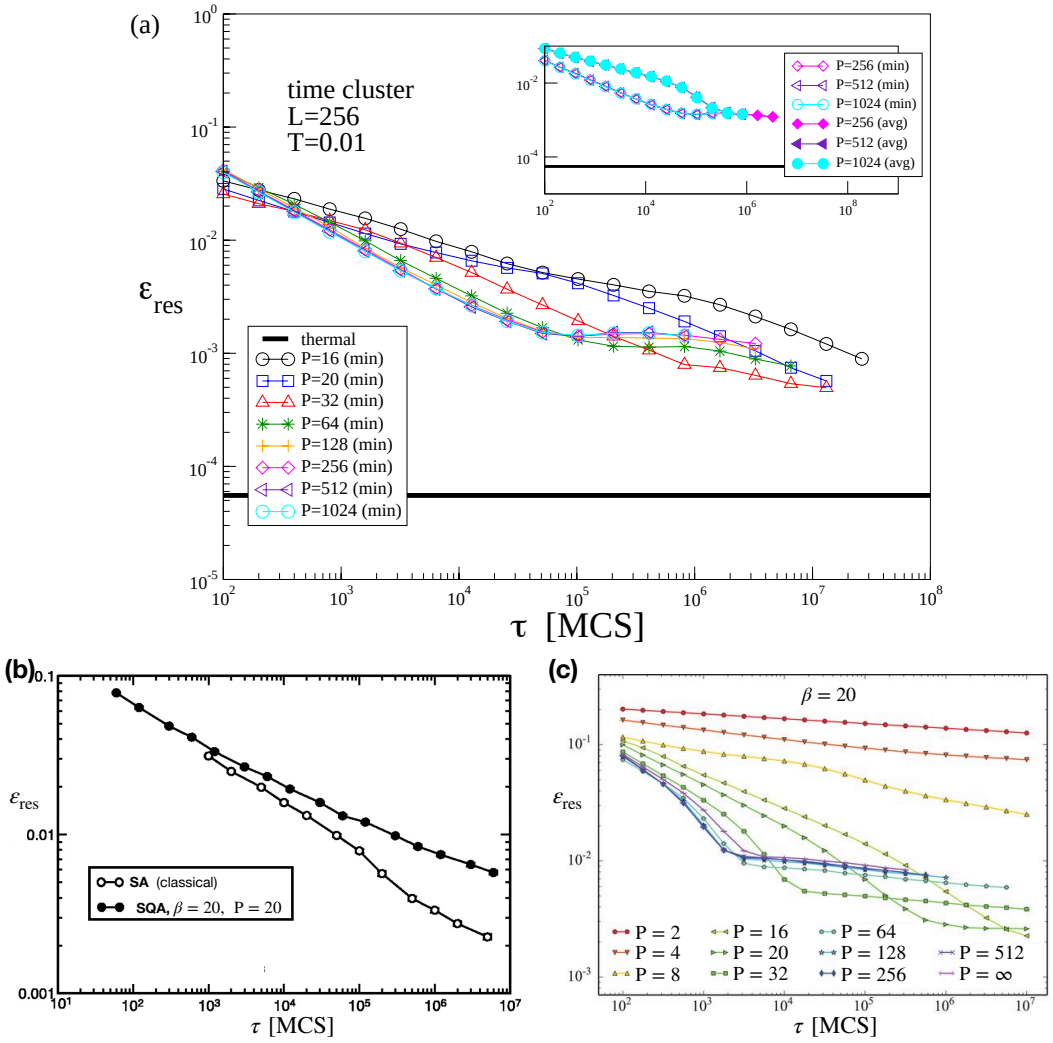

FIGURE 2.4: Residual energy at end of the SQA schedule, with moves local in space, as a function of the annealing time τ. (a) Residual energy SQA with time cluster flips for various values of P and a fixed disorder realization of a disordered Ising Chain. The quantity we plot is the minimum residual energy over all Trotter slices $\varepsilon_{\rm res}^{\rm opt}(\tau)$ (empty symbols), see Eq. (2.15). In the inset we also show the residual energy averaged over all Trotter slices $\varepsilon_{\rm res}^{\rm avg}(\tau)$ (full symbols), see Eq. (2.14). The horizontal thick line denotes the equilibrium thermal value of $\varepsilon_{\rm c}(\Gamma = 0, T = 0.01)$. (b) From from Santoro *et al.* [2002], comparison between residual energies obtained from SA and SQA for a 80×80 disordered Ising model. Here, for SQA, the optimal value P = 20 was chosen to get the best performance. The data corresponds to a single instance of disorder realization. (c) From Heim *et al.* [2015], convergence of SQA towards the continuous limit $\mathrm{P} \to \infty$ for a 80×80 disordered Ising model. The data was obtained by averaging over 1000 disorder realizations.

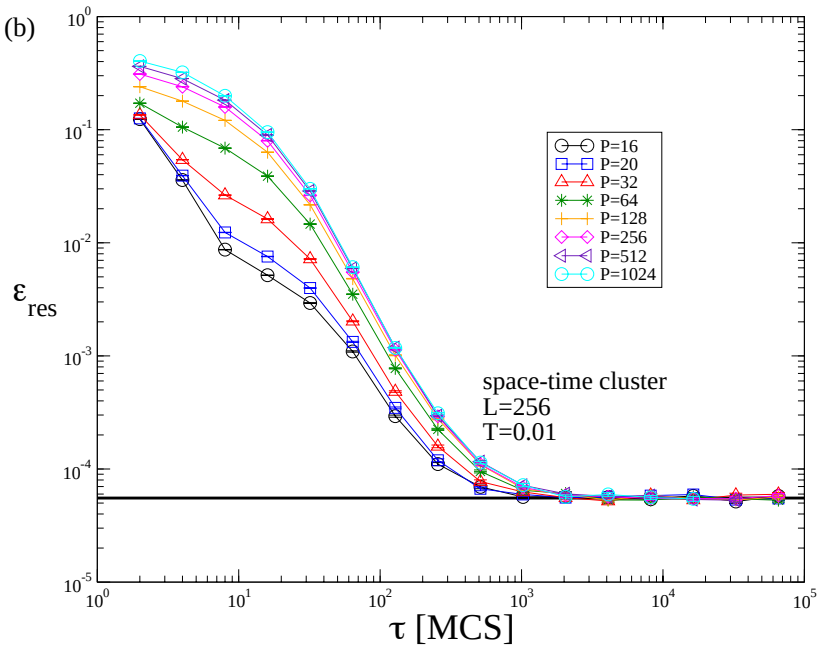

FIGURE 2.5: Residual energy at the end of the SQA schedule with space-time clusters flips as a function of the annealing time. The data is shown for various values of P and a fixed disorder realization. The quantity we plot is the minimum residual energy over all Trotter slices $\varepsilon_{\rm res}^{\rm opt}(\tau)$ (empty symbols), see Eq. (2.15). The horizontal thick line denotes the equilibrium thermal value of $\varepsilon_{\rm c}(\Gamma = 0, T = 0.01)$.

annealing.

However, the question remains: is there any physics that we can learn from the time cluster SQA dynamics in the disordered case? The first tests we have performed consist in monitoring the dynamics of $\epsilon_{\rm res}(t,\tau)$, for a given τ, versus t, both for the QA unitary evolution and the SQA dynamics. Indeed, since each t is univocally associated with a value of $\Gamma(t)$, we can equivalently plot the SQA results, averaged over many repetitions of the Monte Carlo dynamics, versus Γ. Figure 2.6(a) shows the results for three values of τ, with the SQA results denoted by points. Here we find a surprising result, which remains an empirical finding: the SQA with time-cluster moves visits configurations which are essentially equilibrium configurations, but at an effective temperature $T_{\rm eff}(\tau)$, which depends on the total annealing time τ. More precisely, we have verified that the following *Ansatz* for the dynamical residual energy holds:

$$\epsilon_{\rm res}(t,\tau) = \varepsilon_{\rm c}(\Gamma(t), T_{\rm eff}(\tau)) \ , \tag{2.16}$$

where the corresponding equilibrium values of $\varepsilon_{\rm c}(\Gamma, T_{\rm eff})$, with $T_{\rm eff}$ obtained by fitting the numerical points, are shown by dashed lines in Fig. 2.6(a). Even more

remarkably, the same *Ansatz* also holds, on the same disordered instance, for the coherent QA dynamics, performed by integrating through a 4^{th}-order Runge-Kutta algorithm the BdG equations [Caneva *et al.*, 2007; Zanca and Santoro, 2016] for the free-fermion Jordan-Wigner mapping, see Fig. 2.6(b). We might go on and compare the corresponding $T_{\text{eff}}(\tau)$ obtained for the two dynamics. However, since $\varepsilon_{\text{res}}(\tau) = \varepsilon_{\text{c}}(\Gamma = 0, T_{\text{eff}}(\tau))$, according to Eq. (2.16), we can equivalently compare the results obtained for $\varepsilon_{\text{res}}(\tau)$ in the two cases.

Figure 2.7 shows such a comparison. The solid symbols show the SQA results for $\varepsilon_{\text{res}}^{\text{avg}}(\tau)$ in Eq. (2.14), already reported in the inset of Fig. 2.4(a), while the stars show the coherent-QA $\varepsilon_{\text{res}}(\tau)$. Quite evidently, both plots show an intermediate power-law part which, however, shows markedly different power-law exponents in the two cases, $\sim \tau^{-0.9}$ for coherent-QA compared to $\sim \tau^{-0.35}$ for SQA: hence, no linear scaling of the physical against the MC time can ever make the two results consistent. The situation does not improve for large τ, where it is known [Dziarmaga, 2006; Caneva *et al.*, 2007] that, in the thermodynamic limit, the coherent-QA results would display a logarithmic slow-down [Santoro *et al.*, 2002], $\varepsilon_{\text{res}}(\tau) \sim [\log(\gamma\tau)]^{-\xi}$, with $\xi > 2$: the data for $N = 256$ evidently still suffer from finite-size effects which prevent from appreciating such a subtle logarithmic slow-down; nevertheless, they clearly depart from the intermediate-τ data by staying *above* the power-law $\sim \tau^{-0.9}$. On the contrary, the SQA data depart from their intermediate-time power-law $\sim \tau^{-0.35}$ from *below*, but then show a final slowdown of difficult interpretation. The eventual logarithmic behavior that should set in for very large N is, quite evidently, elusive. But the fact that the time-cluster PIMC-SQA dynamics differs from the coherent-QA behavior is quite clear. In principle, decoherence and temperature should be accounted for if a fair comparison of a physical QA dynamics with PIMC-SQA is pursued. We will briefly discuss this issue in the concluding remarks.

2.4 Summary and remarks

We have investigated some aspects of the dynamics behind Simulated Quantum Annealing (SQA), specifically its Path-Integral Monte Carlo (PIMC) implementation, through a detailed analysis of PIMC-SQA on a transverse-field random Ising spin chain, where exact equilibrium and coherent-QA results are easily obtained.

Due to the absence of frustration, we were also able to compare results obtained with two types of Monte Carlo (MC) moves, a local-in-space Swendsen-Wang

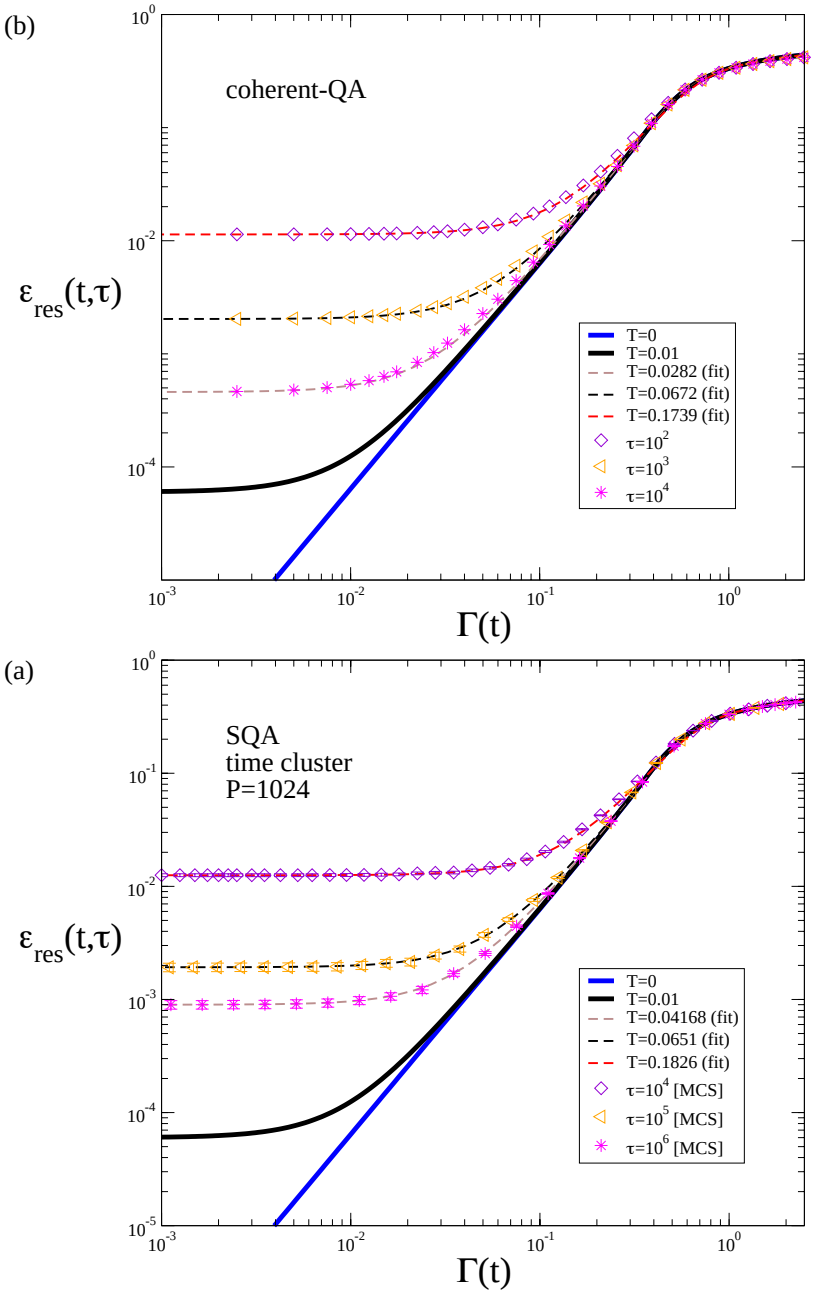

FIGURE 2.6: Test of the dynamical *Ansatz* in Eq. (2.16) for the time-cluster SQA dynamics (a) and the coherent QA dynamics (b), for the same disorder realization of Fig. 2.4(a) and Fig. 2.5 . The numerical data for $\epsilon_{\rm res}(t,\tau)$ versus t are shown by points at the corresponding value for $\Gamma(t)$, while the fits with the equilibrium $\varepsilon_{\rm c}(\Gamma, T_{\rm eff}(\tau))$ are shown by dashed lines. Data for three different values of τ are shown, both for SQA and coherent-QA. For comparison, the exact equilibrium values for $\varepsilon_{\rm c}(\Gamma, T = 0.01)$, at the nominal temperature for the SQA simulation, and $\varepsilon_{\rm c}(\Gamma, T = 0)$, the nominal ground-state value for the coherent-QA dynamics, are also shown by thick solid lines.

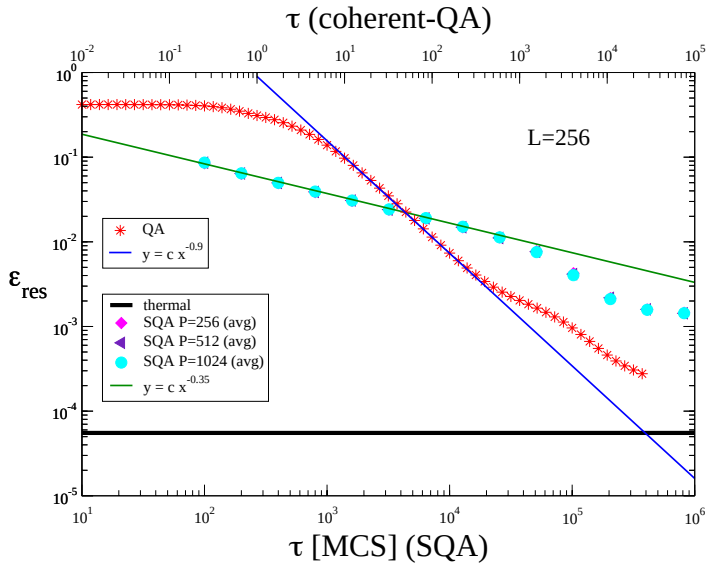

FIGURE 2.7: The residual energy $\varepsilon_{\rm res}(\tau)$ at the end of the annealing for the same disorder realization of Fig. 2.4(a) and Fig. 2.5. The stars show the coherent-QA results vs. τ (upper abscissa axis) while the other symbols refer to the SQA results (the same data are shown in the inset of Fig. 2.4(a)). The equilibrium thermal value $\varepsilon_{\rm c}(\Gamma = 0, T = 0.01)$ is shown by a thick horizontal line. The two straight lines highlight the intermediate power-law regime.

cluster move limited to the imaginary-time direction, with space-time (non-local) SW cluster moves, which provides an extremely fast Monte Carlo dynamics. The results show that the choice of the MC moves is crucial, but that fast non-local cluster moves have nothing to do with any physical dynamics, which is better mimicked by local spin-flip moves.

Concerning the latter more physical choice, we have verified that the expected Kibble-Zurek behavior $\varepsilon_{\rm res}(\tau) \sim \tau^{-1/2}$ is well reproduced by SQA in the ordered case. In the presence of disorder, however, we found out that equilibrium thermodynamical PIMC simulations at finite T show a *sampling problem* emerging, for large P and N, below the critical point $\Gamma < \Gamma_c$ and at low temperatures. The consequences of such a sampling problem on the SQA dynamics are *a priori* not obvious. Interestingly, we found that the time-dependent residual energy $\epsilon_{\rm res}(t, \tau)$ shows features that are shared also by the coherent-QA Schrödinger dynamics, i.e., $\epsilon_{\rm res}(t, \tau)$ is perfectly described by the (instantaneous) equilibrium value of $\varepsilon_{\rm c}(\Gamma(t), T_{\rm eff}(\tau))$ at an effective temperature $T_{\rm eff}(\tau)$ which depends on the annealing time τ. Nevertheless, the SQA results for the residual energy $\varepsilon_{\rm res}(\tau)$ appear to be unrelated, in presence of disorder, with the corresponding coherent-QA results.

Several points still deserve a discussion. One might question the relevance of a comparison of SQA at a finite (low) T with QA results which assume a coherent

Schrödinger evolution in the absence of any external bath. On the practical side, we might add that while a coherent-QA evolution is quite easy to perform here — you just have to integrate $2N \times 2N$ BdG equations for the Jordan-Wigner fermions — the physical dissipative dynamics of a random Ising chain is still a problem which we do not know how to efficiently and reliably tackle. More to the point, however, we can give arguments that are based on our current understanding of the role of dissipation in the QA dynamics of the *ordered* Ising chain [Patanè *et al.*, 2008; Patanè *et al.*, 2009; Arceci *et al.*, 2018]. As indeed shown in Arceci *et al.* [2018], and perhaps easy to argue about, weak dissipation effects have little influence at short and intermediate annealing times τ, which implies that the different power-law behavior displayed in Fig. 2.7 would likely not be influenced by the presence of a bath. For large τ, on the other hand, dissipation tends to drive the system closer to a thermal steady state [Arceci *et al.*, 2018], which might result in a larger residual energy, $\varepsilon_{\rm res}^{\rm diss}(\tau) > \varepsilon_{\rm res}(\tau)$, due to thermal defects generation. Hence, again, it is unlikely that the effect of a weakly coupled thermal bath at temperature T would lead to a closer agreement between SQA and a physical open QA dynamics. The effect of a strongly coupled environment is, currently, beyond our theoretical reach: they constitute a worth pursuing future research direction.

It is worth noticing that the largely accepted viewpoint that the time-continuum limit $\mathrm{P} \to \infty$ is crucial for a comparison with real QA hardware devices. While there is no question to the fact that, as pointed out in Heim *et al.* [2015], a correct quantum mechanical treatment does require the limit $\mathrm{P} \to \infty$, this by itself does not guarantee that the resulting SQA dynamics is *physical*, as we have shown in this Chapter.

The typical "slowdown" that SQA data with $\mathrm{P} \to \infty$ tend to show for large τ should also not necessarily be taken to imply that there is no quantum speedup of any type over classical Simulated Annealing (SA). Indeed, based on theoretical arguments [Santoro *et al.*, 2002], a coherent-QA is expected to show some improvement in the exponent of the logarithmic scaling, $\varepsilon_{\rm res}(\tau) \sim [\log(\gamma\tau)]^{-\xi}$, against competing SA strategies: this improvement has been indeed verified on random Ising chains [Caneva *et al.*, 2007; Suzuki, 2009; Zanca and Santoro, 2016] and on infinitely connected p-spin Ising ferromagnets [Wauters *et al.*, 2017]. Moreover, quite remarkably, non-convex optimization problems are known [Baldassi and Zecchina, 2018] in which SQA, with the $\mathrm{P} \to \infty$ limit properly taken, is much more efficient than its classical SA counterpart.

In the process of studying the $\mathrm{P} \to \infty$ limit, we found that, in the random Ising

chain, the time discretization induced by the Suzuki-Trotter mapping can reduce the density of defect at the end of a linear SQA protocol, thus improving the optimization algorithm. This agrees with previous results obtained for the Ising Spin Glass in two dimensions [Santoro *et al.*, 2002; Heim *et al.*, 2015]. However, since the SQA does not mimic the QA dynamics, it is not clear if a similar behavior is to be expected for a coherent QA dynamics in a physical device. In the next Chapter, we address this issue by studying in detail the effect of a Suzuki-Trotter discretization in linear coherent-QA protocols.

Chapter 3

Optimal working point for linear schedules in digitized QA

We present a study of the *digitized*-QA (dQA) protocol proposed by Barends *et al.* [2016]. Our analysis, performed on the benchmark case of a transverse Ising chain problem, shows that the algorithm has a well defined *optimal working point* for the annealing time $\tau_{\mathrm{P}}^{\mathrm{opt}}$ — scaling as $\tau_{\mathrm{P}}^{\mathrm{opt}} \sim \mathrm{P}$, where P is the number of digital Trotter steps — beyond which the residual energy error shoots-up towards the value characteristic of the maximally disordered state. We present an analytical analysis for the translationally invariant transverse Ising chain, but our numerical evidence suggests that this scenario is more general, surviving, for instance, the presence of disorder.

3.1 Introduction

The QA approach is intrinsically *analog*: it requires constructing the appropriate hardware for a given problem Hamiltonian $\widehat{H}_z$ and a continuous tuning of the schedule parameter$s(t)$ during runtime. A recent experimental work by the group of Martinis [Barends *et al.*, 2016] has advocated a *digitized*-QA, in essence a discrete-time Trotter-Suzuki annealing dynamics. In the simplest cases, one constructs an approximate state through a depth-P digital quantum circuit by splitting the QA Hamiltonian into two non-commuting pieces $\widehat{H}_z$ and $\widehat{H}_x$ (involving, say, $\hat{\sigma}^z$ and $\hat{\sigma}^x$ Pauli operators):

$$|\psi_{\mathrm{P}}\rangle = \mathrm{e}^{-i\beta_{\mathrm{P}}\widehat{H}_x}\mathrm{e}^{-i\gamma_{\mathrm{P}}\widehat{H}_z}\cdots\mathrm{e}^{-i\beta_1\widehat{H}_x}\mathrm{e}^{-i\gamma_1\widehat{H}_z}|\psi_0\rangle\ . \tag{3.1}$$

The *digitized*-QA parameters $\boldsymbol{\beta} = (\beta_1 \cdots \beta_{\mathrm{P}})$ and $\boldsymbol{\gamma} = (\gamma_1 \cdots \gamma_{\mathrm{P}})$ are obtained by applying appropriate Suzuki-Trotter splitting to the time discretized Schrödinger evolution operator (see Eq. (1.22)). One great advantage of such a *digitized*-QA approach is that the different unitaries can be "constructed" using universal gates, implementing, for instance, $\hat{\sigma}_i^z \hat{\sigma}_j^z$ and $\hat{\sigma}_i^x \hat{\sigma}_j^x$ couplings, thus allowing for the inclusion of *non-stoquastic* [Bravyi *et al.*, 2008] terms, generally hard to simulate on a classical computer [Troyer and Wiese, 2005]. A second advantage of a digital implementation is, as argued in Barends *et al.* [2016], that the system can in principle be made fully fault-tolerant [Bravyi and Kitaev, 1998; Fowler *et al.*, 2012].

The dQA form of the state in Eq. (3.1) bears an appealing similarity with a hybrid quantum-classical approach known as QAOA, introduced by Farhi *et al.* [Farhi *et al.*, 2014]. The non-trivial difference with dQA is that in QAOA the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are assumed to be *variational parameters* of the state $|\psi_{\mathrm{P}}(\boldsymbol{\beta}, \boldsymbol{\gamma})\rangle$, which should be appropriately optimized, via some classical algorithm.

Even without optimization of the parameters, a plain linear-schedule dQA still requires a careful choice of the number of Trotter steps P, for a given total annealing time τ. This is precisely the issue we analyze in this Chapter. By considering the usual benchmark case of a transverse Ising chain, we show that for any fixed finite P, a linear-schedule dQA has a clear *optimal working point* annealing time $\tau_{\mathrm{P}}^{\mathrm{opt}}$ — scaling as $\tau_{\mathrm{P}}^{\mathrm{opt}} \sim \mathrm{P}$ — beyond which the residual energy error shoots-up towards the value characteristic of the maximally disordered state. Interestingly, we show that the time-discretization and Trotter errors implied by dQA always lead to a residual energy which is *larger* than the corresponding continuous-time QA, at variance with what we saw in Chapter 2. The scaling behavior of linear-dQA at the optimal point, however, is identical with the standard Kibble-Zurek scaling [Kibble, 1976; Zurek, 1985; Polkovnikov *et al.*, 2011] predicted for continuous-time QA with a linear schedule [Zurek *et al.*, 2005; Dziarmaga, 2005].

The rest of the Chapter is organized as follows. In Sec. 3.2 we briefly recall the model and the different variants of QA that we consider. In Sec. 3.3 we present our results for *digitized*-QA with a linear annealing schedule for a translationally invariant quantum Ising chain, discussing the presence of the optimal working point $\tau_{\mathrm{P}}^{\mathrm{opt}}$ for different values of the Trotter time-steps P. In Sec. 3.4 we show the effects of time-discretization compared to Trotter digitalization errors. In Sec. 3.5 we discuss the effects of disorder. Section 3.6 finally, presents our summary and a final discussion of relevant points.

3.2 Model and methods

The class of models considered in Barends *et al.* [2016] is described by the usual QA interpolating Hamiltonian (see Eq. (1.1)):

$$\widehat{\mathrm{H}}(s) = s\widehat{H}_{\mathrm{T}} + (1-s)\widehat{H}_{\mathrm{D}} \ , \tag{3.2}$$

where $s \in [0,1]$ is a rescaled time, $\widehat{H}_{\mathrm{D}} = -\Gamma \sum_{j=1}^{N} \hat{\sigma}_j^x$ is the usual transverse field driving term, and $\widehat{H}_{\mathrm{T}}$ is the target Hamiltonian of which we would like to compute the ground state:

$$\widehat{H}_{\mathrm{T}} = -\sum_{j=1}^{N} \left[J_j^z \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z + J_j^x \hat{\sigma}_j^x \hat{\sigma}_{j+1}^x + B_j^z \hat{\sigma}_j^z + B_j^x \hat{\sigma}_j^x \right] . \tag{3.3}$$

Here $\left(\hat{\sigma}_j^x, \hat{\sigma}_j^y, \hat{\sigma}_j^z\right)$ are Pauli matrices on the j^{th} site, and N the number of sites. The transverse field Ising case is recovered for $J_j^x = 0$, which can be exactly solved through a Jordan-Wigner mapping [Lieb *et al.*, 1961; Young, 1997] if we further set also $B_j^z = 0$. The initial state is the ground state of $\widehat{H}_{\mathrm{D}}$, namely $|\psi_0\rangle = |+\rangle^{\otimes N}$ (see Eq. (1.7)).

A brief introduction to *digitized*-QA was provided in Chapter 1. However, since the effects of time-discretization and full digitalization are central issues here, we re-discuss them here in more detail, confining the discussion to a linear schedule. The exact continuous-time QA evolution is given by the time-ordered exponential evolution operator

$$\widehat{U}_{\mathrm{QA}}(\tau, 0) = \mathrm{Texp}\left(-\frac{i}{\hbar} \int_0^\tau \mathrm{d}t \ \widehat{\mathrm{H}}(s(t)) \right) \ . \tag{3.4}$$

The first step towards full digitalization consists in a time-discretization: For the simple case of a linear schedule QA where $s(t) = t/\tau$, with τ the total annealing time, we divide τ into P time-intervals $\Delta t = \tau/\mathrm{P}$, and write:

$$\widehat{U}_{\mathrm{QA}}(\tau, 0) \Longrightarrow \widehat{U}_{\mathrm{step}} = \mathrm{e}^{-\frac{i\Delta t}{\hbar}\widehat{\mathrm{H}}(s_{\mathrm{P}})} \cdots \mathrm{e}^{-\frac{i\Delta t}{\hbar}\widehat{\mathrm{H}}(s_1)} \ , \tag{3.5}$$

where $s_m = m/\mathrm{P}$, with $m = 1, \cdots, \mathrm{P}$. To achieve a full digitalization, one must express $\mathrm{e}^{-\frac{i\Delta t}{\hbar}\widehat{\mathrm{H}}(s_m)}$ in terms of simple quantum gates [Barends *et al.*, 2016]. Let us denote by $\widehat{\mathrm{H}}_x(s) = s\widehat{H}_{\mathrm{T}}^x + (1-s)\widehat{H}_{\mathrm{D}}$ and $\widehat{\mathrm{H}}_z(s) = s\widehat{H}_{\mathrm{T}}^z$, the terms in $\widehat{\mathrm{H}}(s)$ containing $\hat{\sigma}^x$ and $\hat{\sigma}^z$ respectively. The simplest possible Trotter splitting of $\widehat{\mathrm{H}}(s) =$

$\widehat{\mathrm{H}}_x(s) + \widehat{\mathrm{H}}_z(s)$ is the 1st-order one:

$$\mathrm{e}^{-i\epsilon\widehat{\mathrm{H}}(s)} \simeq \mathrm{e}^{-i\epsilon\widehat{\mathrm{H}}_x(s)}\mathrm{e}^{-i\epsilon\widehat{\mathrm{H}}_z(s)} + O(\epsilon^2) \,, \tag{3.6}$$

which can be easily improved by a symmetrized 2nd-order splitting:

$$\mathrm{e}^{-i\epsilon\widehat{\mathrm{H}}(s)} \simeq \mathrm{e}^{-i\frac{\epsilon}{2}\widehat{\mathrm{H}}_x(s)}\mathrm{e}^{-i\epsilon\widehat{\mathrm{H}}_z(s)}\mathrm{e}^{-i\frac{\epsilon}{2}\widehat{\mathrm{H}}_x(s)} + O(\epsilon^3) \,. \tag{3.7}$$

For ease of notation, and to concentrate on the cases we will actually consider in the numerics, we set $J_j^x = B_j^x = B_j^z = 0$. Then, the target Hamiltonian $\widehat{H}_z$ describes a random Ising chain introduced in Chapter 2 (see Eq. (2.1) and Eq. (2.2)). Namely, we have

$$\widehat{H}_z \;=\; -\sum_{j=1}^{N} J_j^z \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z \tag{3.8}$$

$$\widehat{H}_x \;=\; \widehat{H}_{\mathrm{D}} = -\Gamma \sum_{j=1}^{N} \hat{\sigma}_j^x \tag{3.9}$$

where, Periodic Boundary Conditions[1] (PBC) are implicitly assumed, so that $\hat{\sigma}_{N+1}^{x,y,z} \stackrel{\mathrm{def}}{=} \hat{\sigma}_1^{x,y,z}$. Then the fully *digitized*-QA evolution operator can be written as:

$$\widehat{U}_{\mathrm{digit}} = \mathrm{e}^{-i\beta_{\mathrm{P}}\widehat{H}_x}\mathrm{e}^{-i\gamma_{\mathrm{P}}\widehat{H}_z}\cdots\mathrm{e}^{-i\beta_1\widehat{H}_x}\mathrm{e}^{-i\gamma_1\widehat{H}_z} \,, \tag{3.10}$$

where

$$\gamma_m = s_m\frac{\Delta t}{\hbar} \quad \text{with} \quad s_m = \frac{m}{\mathrm{P}} \,, \tag{3.11}$$

while

$$\beta_m = (1 - s_m)\frac{\Delta t}{\hbar} \qquad (1^{\mathrm{st}}\text{-order Trotter}) \tag{3.12a}$$

$$\beta_m = \big(1 - \frac{s_m + s_{m+1}}{2}\big)\frac{\Delta t}{\hbar} \qquad (2^{\mathrm{nd}}\text{-order Trotter}) \tag{3.12b}$$

with the proviso that we should take $\beta_{\mathrm{P}} = 0$ and neglect an irrelevant phase factor $\mathrm{e}^{i\Gamma N\Delta t/(2\hbar)}$ appearing in the symmetric 2nd-order splitting.

We observe that $\widehat{U}_{\mathrm{step}}$ introduces a *time-discretization error*, while $\widehat{U}_{\mathrm{digit}}$ introduces an additional *digital (or Trotter) error* [Barends *et al.*, 2016] associated with the non commutativity of the quantum operators appearing in $\widehat{U}_{\mathrm{step}}$.

[1] Note that in Chapter 2 OBC were assumed instead of PBC.

To assess the quality of the annealing, we still monitor the density of defects (previously introduced in Eq. (2.3)) created over the ferromagnetic classical Ising ground state[2]:

$$\rho_{\rm def}(\tau) = \frac{1}{2N}\sum_{j=1}^{N}\langle\psi(\tau)|(1-\hat{\sigma}_j^z\hat{\sigma}_{j+1}^z)|\psi(\tau)\rangle\ , \tag{3.13}$$

where the expectation value is taken over the final state at time $t=\tau$, $|\psi(\tau)\rangle = \widehat{U}(\tau)|\psi_0\rangle$, calculated with different evolution operators. More precisely, we will compare:

1. The continuous-time QA with a linear-schedule $s(t)=t/\tau$, whose evolution operator is given by Eq (3.4);
2. The corresponding fully *digitized*-QA (dQA) with $\widehat{U}(\tau)=\widehat{U}_{\rm digit}$ in Eq. (3.10) with a symmetric Trotter splitting, see Eqs. (3.11) and (3.12b);
3. The step-discretized QA evolution with $\widehat{U}(\tau)=\widehat{U}_{\rm step}$ in Eq. (3.5), to discriminate digital errors from time-discretization errors.

3.3 Results for digitized-QA on Ising chain

Particularly simple is the case of the translationally invariant transverse Ising model, $J_j^z = J$, with periodic boundary conditions (PBC), which reduces to an assembly of 2×2 independent problems labeled by the momentum k, depending on the fermionic parity [Dziarmaga, 2005]. A Jordan-Wigner transformation [Jordan and Wigner, 1928; Lieb *et al.*, 1961], $\hat{\sigma}_j^x = 1-2\hat{c}_j^\dagger\hat{c}_j$, $\hat{\sigma}_j^z = -(\hat{c}_j+\hat{c}_j^\dagger)\exp\left(-i\pi\sum_{l=1}^{j-1}\hat{c}_l^\dagger\hat{c}_l\right)$, maps the spin system to free spinless fermions on a lattice, where $\hat{c}_j^\dagger$ and $\hat{c}_j$ respectively create and annihilate a fermion at site j. A Fourier transform can then be used to decompose the system into a set of decoupled two-level systems. The transformation is standard, see App. A. The final result can be cast in the following form:

$$\widehat{H}_x = 2\Gamma\sum_{k>0}^{\rm ABC}\left(\hat{c}_k^\dagger\hat{c}_k - \hat{c}_{-k}\hat{c}_{-k}^\dagger\right) \tag{3.14}$$

and

$$\widehat{H}_z = -2J\sum_{k>0}^{\rm ABC}\left(\cos k\left(\hat{c}_k^\dagger\hat{c}_k-\hat{c}_{-k}\hat{c}_{-k}^\dagger\right)+\sin k\left(\hat{c}_k^\dagger\hat{c}_{-k}^\dagger+\hat{c}_{-k}\hat{c}_k\right)\right)\ , \tag{3.15}$$

[2] Here, PBC are implicitly also assumed in the definition of $\rho_{\rm def}(\tau)$.

where the sum over k runs over the anti-periodic boundary conditions (ABC) positive wave-vectors given by $k = \pi(2n-1)/N$, with $n = 1, 2, \ldots, N/2$, corresponding to the even-fermion-parity sector [Dziarmaga, 2005]. It is convenient to rewrite these Hamiltonians in terms of $N/2$ two-level-systems with effective Pauli matrices $\hat{\boldsymbol{\tau}}_k = (\hat{\tau}_k^x, \hat{\tau}_k^y, \hat{\tau}_k^z)^T$ labelled by the $N/2$ independent k-vectors. With the convention that $|\uparrow_k\rangle = |0\rangle$ and $|\downarrow_k\rangle = \hat{c}_k^\dagger \hat{c}_{-k}^\dagger |0\rangle$, it is simple to check that the Hamiltonian terms translate as follows (see App.A):

$$\widehat{H}_x = -2\Gamma \sum_{k>0}^{\text{ABC}} \hat{\tau}_k^z = -2\Gamma \sum_{k>0}^{\text{ABC}} \hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{\tau}}_k \tag{3.16}$$

and

$$\widehat{H}_z = 2J \sum_{k>0}^{\text{ABC}} \left(\cos k\, \hat{\tau}_k^z - \sin k\, \hat{\tau}_k^x \right) = -2J \sum_{k>0}^{\text{ABC}} \hat{\boldsymbol{b}}_k \cdot \hat{\boldsymbol{\tau}}_k \ , \tag{3.17}$$

where we have introduced two unit vectors

$$\hat{\boldsymbol{z}} = (0, 0, 1)^T \tag{3.18}$$

$$\hat{\boldsymbol{b}}_k = (\sin k, 0, -\cos k)^T \ . \tag{3.19}$$

Observe that, with the previous definitions, the density of defects in Eq. (3.13) simply reads:

$$\rho_{\text{def}}(\tau) = \frac{1}{2} - \frac{1}{N} \sum_{k>0}^{\text{ABC}} \hat{\boldsymbol{b}}_k \cdot \langle \psi_0 | \widehat{U}^\dagger(\tau) \hat{\boldsymbol{\tau}}_k \widehat{U}(\tau) | \psi_0 \rangle \ . \tag{3.20}$$

We start by studying how the defect density depends on the total annealing time τ for a *digitized*-QA with a fixed number P of quantum gates. Henceforth we fix $J = \Gamma$ and measure times in units of $\hbar/J$, which is equivalent to setting $\hbar = 1$ and $J = 1$. In Fig. 3.1 we plot the density of defects $\rho_{\text{def}}(\tau)$ versus annealing time τ for a translationally invariant Ising chain with $N = 512$ sites. The solid line shows the results of a continuous-time linear-schedule QA — obtained by numerically solving the time-dependent Schrödinger equation —, clearly displaying a Kibble-Zurek [Kibble, 1976; Zurek, 1985; Polkovnikov *et al.*, 2011] (KZ) scaling behaviour [Zurek *et al.*, 2005; Dziarmaga, 2005]. The different dotted lines show the results obtained with $\widehat{U}_{\text{digit}}$, labelled "linear-dQA", for three values of $\text{P} = 4, 32, 128$. As expected, for short and intermediate annealing times ($\tau < \text{P}$) time-discritization and digital errors are negligible. When τ is increased further, the digital protocols show a sharp *optimal working point* $\tau_{\text{P}}^{\text{opt}}$, which depends on P roughly as

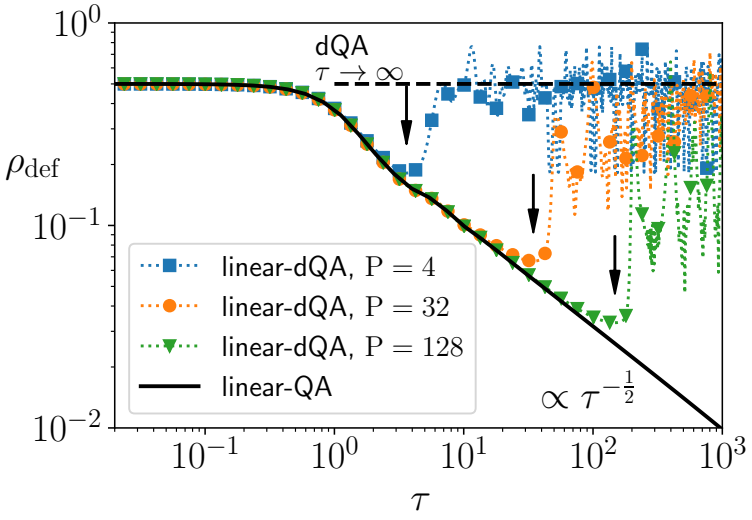

FIGURE 3.1: Density of defects $\rho_{\rm def}(\tau)$ versus annealing time τ for an translationally invariant quantum Ising chain with $N = 512$ sites and different fixed values of P, corresponding to Trotter time-steps of $\Delta t = \tau/\mathrm{P}$. The vertical arrows indicate the optimal working points $\tau_{\rm P}^{\rm opt}$.

$\tau_{\rm opt}(\mathrm{P}) \propto \mathrm{P}$, before $\rho_{\rm def}(\tau)$ shoots-up towards a rather irregular behaviour for $\tau \to \infty$: we will analytically show below that $\rho_{\rm def}(\tau)$ oscillates around the average value $\rho_{\infty}^{\rm digit} = 1/2$, corresponding to the maximally disordered state. The optimal working point is particularly crucial in the dQA protocol, as increasing τ beyond $\tau_{\rm opt}$ would completely spoil the result and waste resources. Incidentally, the large-τ irregular behavior of the *digitized*-QA data is not a finite-N effect: as we explain in Chapter 4, the thermodynamic limit is effectively reached as soon as $N \geq 2\mathrm{P} + 2$.

To understand the $\tau \to \infty$ regime at fixed P, we start by observing that a good way to extract the asymptotic behaviour for $\rho_{\rm def}(\tau \to \infty)$ is to calculate an *infinite-time average* of $\rho_{\rm def}(\tau)$:

$$\overline{\rho}_{\rm def} = \lim_{\mathrm{T}\to\infty} \frac{1}{\mathrm{T}} \int_0^{\mathrm{T}} \mathrm{d}\tau\, \rho_{\rm def}(\tau)\ . \tag{3.21}$$

For a translationally invariant chain in the thermodynamic limit, where the different k-modes become a continuum, exchanging the time-average integral with the k-integral, see Eqs. (3.20) and (3.21), we get:

$$\overline{\rho}_{\rm def} = \frac{1}{2} - \int_0^{\pi} \frac{\mathrm{d}k}{2\pi}\, \hat{\boldsymbol{b}}_k \cdot \overline{\langle \hat{\boldsymbol{\tau}}_k \rangle_\tau} \tag{3.22}$$

where $\overline{\langle\hat{\boldsymbol{\tau}}_k\rangle_\tau}$ is the infinite-time average of:

$$\langle\hat{\boldsymbol{\tau}}_k\rangle_\tau = \langle\psi_k(0)|\widehat{U}_k^\dagger(\tau)\hat{\boldsymbol{\tau}}_k\widehat{U}_k(\tau)|\psi_k(0)\rangle \ . \tag{3.23}$$

The form of the unitary $\widehat{U}_k(\tau)$ depends on the QA protocol used. The *digitized*-QA case is particularly simple: one can express the action on $\hat{\boldsymbol{\tau}}_k$ of the relevant 2×2 unitary operators — $\mathrm{e}^{i2\beta_m\hat{\boldsymbol{z}}\cdot\hat{\boldsymbol{\tau}}_k}$ and $\mathrm{e}^{i2\gamma_m\hat{\boldsymbol{b}}_k\cdot\hat{\boldsymbol{\tau}}_k}$, in alternation, see Eqs. (3.10), (3.16) and (3.17) — through a product of 3×3 rotation matrices $\mathcal{R}$, by repeatedly using the following Pauli matrix identity:

$$e^{-i\frac{\theta}{2}\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{\tau}}}\hat{\boldsymbol{\tau}}e^{i\frac{\theta}{2}\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{\tau}}} = \mathcal{R}_{\hat{\boldsymbol{n}}}(\theta)\hat{\boldsymbol{\tau}} \ , \tag{3.24}$$

where $\hat{\boldsymbol{n}}$ is the axis of rotation, and θ the rotation angle. As a result of that, for the *digitized*-QA case we end up writing:

$$\overline{\langle\hat{\boldsymbol{\tau}}_k\rangle_\tau} = \overline{\left[\prod_{m=1}^{\leftarrow \mathrm{P}} \mathcal{R}_{\hat{\boldsymbol{z}}}(4\beta_m)\mathcal{R}_{\hat{\boldsymbol{b}}_k}(4\gamma_m)\right]\hat{\boldsymbol{z}}} \tag{3.25}$$

where $\prod^{\leftarrow \mathrm{P}}$ denotes a time-ordered product, and we explicitly used that for the initial state $\langle\psi_k(0)|\hat{\boldsymbol{\tau}}_k|\psi_k(0)\rangle = \hat{\boldsymbol{z}}$ holds. Full details of this analysis can be found in App. B. Concerning the infinite-τ average of $\langle\hat{\boldsymbol{\tau}}_k\rangle_\tau$, we observe that *if* the different rotation matrices are *uncorrelated*, then we can transform the complicated τ-average of a product into a much simpler product of τ-averages. It turns out that the fully digital case $\widehat{U}_{\mathrm{digit}}$ behaves precisely in this way, and one can also show (see App. B):

$$\int_0^\pi \frac{\mathrm{d}k}{2\pi}\,\hat{\boldsymbol{b}}_k\cdot\left[\prod_{m=1}^{\leftarrow \mathrm{P}} \overline{\mathcal{R}_{\hat{\boldsymbol{z}}}(4\beta_m)}\;\overline{\mathcal{R}_{\hat{\boldsymbol{b}}_k}(4\gamma_m)}\right]\hat{\boldsymbol{z}} = 0 \ , \tag{3.26}$$

which implies that $\overline{p}_{\mathrm{def}}^{\mathrm{digit}} = \frac{1}{2}$.

So far we considered *digitized*-QA at fixed P. One might ask what happens if one considers results for increasing P at constant $\Delta t = \tau/\mathrm{P}$, corresponding, in the experiment, to applying gates of fixed time-duration. Figure 3.2 shows the numerical result we obtained. It is worth noting that the best results are obtained when $\Delta t \approx 1$ (in units of $\hbar/J$), in a way that is consistent with the optimal working point shown in Fig. (3.1), and with the Kibble-Zurek scaling exponent [Zurek *et al.*, 2005; Dziarmaga, 2005]. For $\Delta t \ll 1$ the Trotter error is negligibly

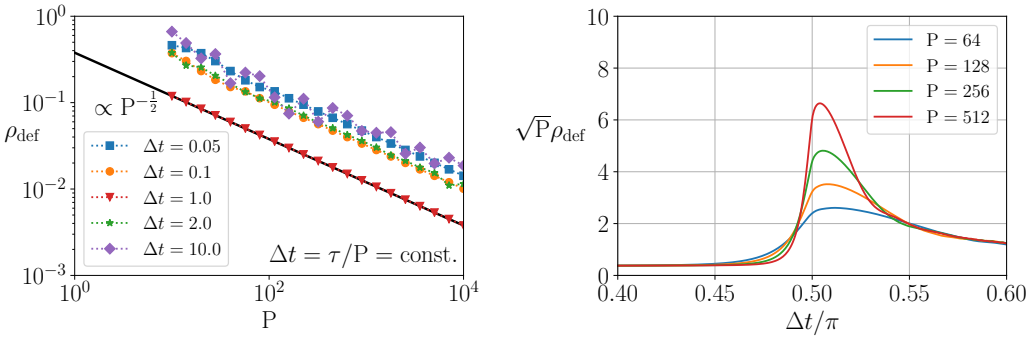

FIGURE 3.2: (top) Density of defects $\rho_{\rm def}(\tau)$ after annealing at several fixed $\Delta t = \tau/\mathrm{P}$ as a function of P for an ordered Ising chain with $N = 1024$. (bottom) Density of defects $\rho_{\rm def}(\tau)$ after annealing at several fixed P as a function of $\Delta t = \tau/\mathrm{P}$ for an translationally invariant Ising chain with $N = 1024$.

small, but we are wasting resources. For $\Delta t \sim 1$ the Trotter error is not small, but the *digitized*-QA dynamics is very effective and indeed optimal. For $\Delta t \gtrsim \pi/2$, as shown in Fig. 3.2(b), there is a sudden increase in the defect density, reflecting the fact that the digitized-QA dynamics is no longer adiabatic[3]: this completely spoils the quality of the results.

3.4 Time-discretization versus digitalization

We now compare the result of a full digitalization with those obtained by a time-discretization of the evolution operator in P time steps of $\Delta t = \tau/\mathrm{P}$, which we will denote as *step*-QA, see Eq. (3.5), with the same linear schedule $s(t) = t/\tau$. Figure 3.3 shows the comparison between *digitized*-QA and *step*-QA for the usual translationally invariant Ising chain. Notice how the behaviour is very similar for small $\tau < \mathrm{P}$ — indeed almost indistinguishable from the continuous-time QA result. For larger τ, however, the behaviour is radically different: $\rho_{\rm def}(\tau)$ computed with $\widehat{U}_{\rm step}$ saturates to a finite plateau value which decreases as P is increased, at variance with the *digitized*-QA results: time-discretization and digital errors are clearly distinguishable in that regime. It turns out that an analytical determination of the plateau value is still possible even in the *step*-QA case, but the algebra is considerably more involved, because of correlations between the different rotation matrices applied to $\hat{\boldsymbol{\tau}}$. Without entering into details, reported in App. B, we just mention that the calculation of the infinite time average in Eq. (3.22) is reduced, for a given value of P, to a contour integral over the unit

[3] We postpone a discussion of adiabaticity in digitized protocols to Chapter 4. See also [Dranov *et al.*, 1998; Mbeng *et al.*, 2019].

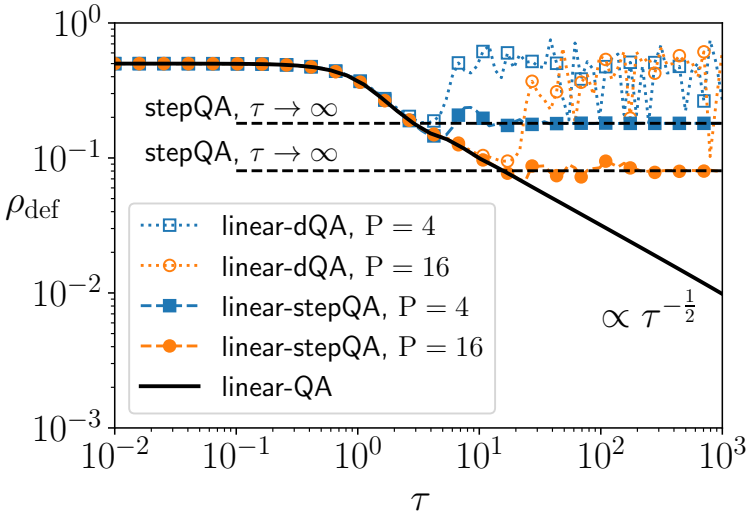

FIGURE 3.3: Density of defects $\rho_{\rm def}(\tau)$ versus annealing time τ for a translationally invariant quantum Ising chain. We compare *digitized*-QA with *step*-QA results, obtained by a time-discretization of the evolution operator as in Eq. (3.5). Here P = 4 and 16 while $N = 1024$. The horizontal solid lines are the analytical predictions for the infinite-time averages in the *step*-QA case, $\overline{\rho}^{\rm step}_{\rm def,P=4} = \frac{13}{72}$ and $\overline{\rho}^{\rm step}_{\rm def,P=16}$.

circle C in the complex plane of a rational function

$$\overline{\rho}^{\rm step}_{\rm def,P} = \frac{1}{2} + \frac{1}{2\pi i} \oint_C {\rm d}z \frac{f_{\rm P}(z)}{g_{\rm P}(z)} \,, \tag{3.27}$$

where $f_{\rm P}(z)$ and $g_{\rm P}(z)$ are polynomials of the complex variable z. The integrand (for even P) has P/2 2$^{\rm nd}$-order poles given by the roots of the polynomial $g_{\rm P}(z)$ inside the unit circle C, located at $z_m = \frac{m}{{\rm P}-m}$ with $m = 0, 1, \cdots, ({\rm P}/2 - 1)$. By calculating explicitly the sum of the residues we find that, for instance, $\overline{\rho}^{\rm step}_{\rm def,P=2} = \frac{1}{4}$, $\overline{\rho}^{\rm step}_{\rm def,P=4} = \frac{13}{72}$, $\overline{\rho}^{\rm step}_{\rm def,P=6} = \frac{18631}{128000}$, while higher P lead to large fractions which we can calculate with Mathematica. As shown in Fig. 3.3, our analytical prediction matches very well the numerical simulations.

3.5 The effect of disorder

The main features we have found so far are not limited to a translationally invariant Ising model. In principle, an analysis based on unitary operators can still pursued in the presence of disorder, as long as one can apply a Jordan-Wigner transformation obtaining a quadratic fermionic Hamiltonian. The matrices are now $2N \times 2N$ rather than 2×2. We have not carried out such an analytical

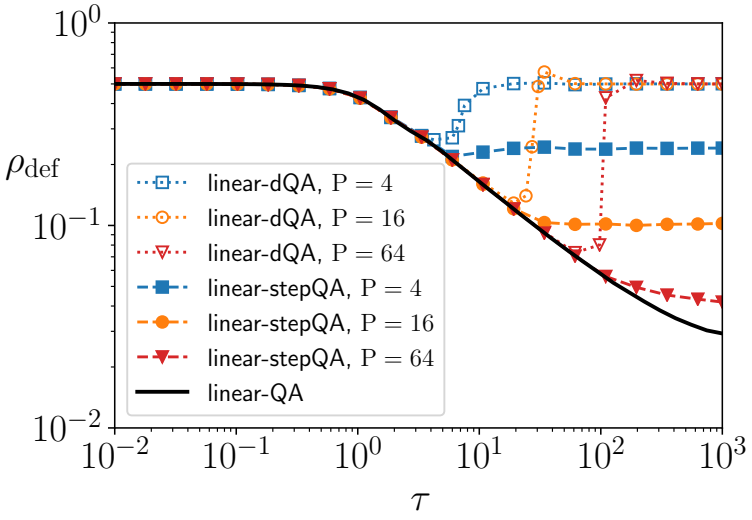

FIGURE 3.4: Density of defects $\rho_{\rm def}(\tau)$ versus annealing time τ for a random quantum Ising chain of length $N = 128$, averaged over 1000 realizations of disorder with $J_j^z \in (0, 1]$.

calculation. However, as we did in Chapter 2, one can calculate numerically the annealing results for $\rho_{\rm def}(\tau)$ with the various QA protocols, reported in Fig. 3.4 for a disordered chain with $N = 128$. Here the couplings are taken to be $J_j^z \in (0, 1]$, and the results are averaged over 1000 realizations of the disorder. The overall features observed are fully consistent with the ordered case results. In particular, there is a sharp optimal working point for the *digitized*-QA case, where once again $\tau_{\rm P}^{\rm opt} \propto {\rm P}$.

3.6 Summary and remarks

Summarizing, we have analyzed the effects of Trotter error in *digitized*-QA, with a linear annealing schedule, for the benchmark case of a transverse field Ising chain, highlighting the presence of a sharp optimal working point with $\tau_{\rm P}^{\rm opt} \propto {\rm P}$, which should be taken care of, to avoid wasting resources or even spoiling the annealing results with large digital errors. The results we have obtained are consistent with the clear minima observed in the experiment, see in particular Fig. 3.5 from Barends *et al.* [2016]. Interestingly, at the optimal working point the scaling behavior is precisely consistent with the Kibble-Zurek behavior [Zurek *et al.*, 2005; Dziarmaga, 2005] seen for continuous-time QA with a linear schedule $s(t) = t/\tau$. We have also analyzed the effect of time-discretization without Trotter

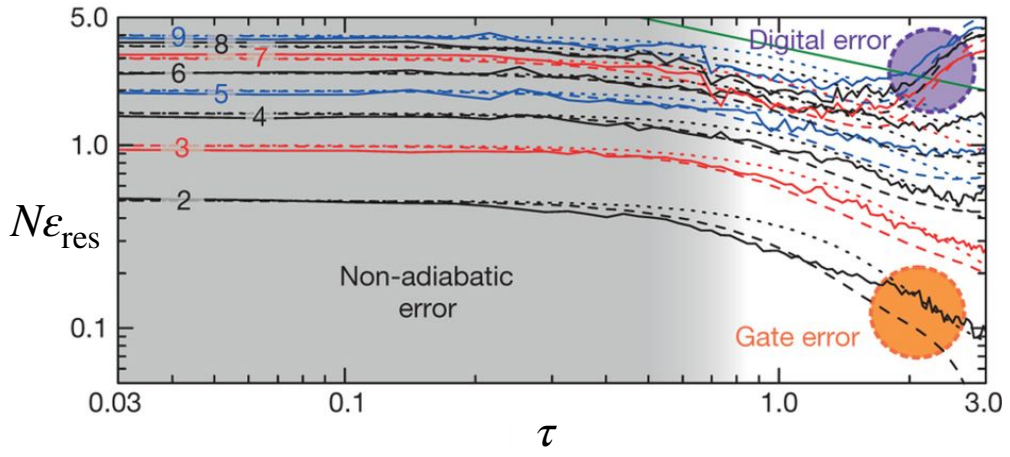

FIGURE 3.5: Residual energy at the end of a *digitized*-QA schedule as a function of the annealing time. The data shown were obtained in the experiment reported in Barends *et al.* [2016], where the *digitized*-QA algorithm is run on a system of superconducting qubits. The experiment simulates a translationally invariant quantum Ising chain for different system sizes $N = 2, 3, \cdots, 9$ (as indicated by the color-coded numerals). The green solid line shows the KZ scaling $\epsilon_{\rm res} \sim \tau^{-\frac{1}{2}}$. Figure taken from Barends *et al.* [2016].

splitting, showing that the digital error behaves in a drastically different way for long annealing times.

We also observed that, for coherent *digital*-QA, digital errors are seen to lead to a density of defects which is always *larger* than that for continuous-time QA. This is at variance with what we saw for SQA in Chapter 2, where digital errors can *lower* the density of defects for intermediate annealing times. It would be extremely interesting to know what digitalization would do in a context of open quantum system dynamics, where the effect of the environment on the QA dynamics is duly accounted for. This is an issue worth investigating, which we leave for future studies.

We conclude the Chapter by noting that, although choosing the optimal Δt leads to a computational advantage, the complexity (or scaling) of the algorithm remains unchanged. Indeed, as we anticipated in Chapter 1, to drastically improve the performance of QA, we should consider more general non-linear schedules $s(t)$. Therefore, in the next Chapter, we go beyond the simple optimization of *linear digitized*-QA and discuss how to find optimal *non-linear schedules*.

Chapter 4

Making digitized Quantum Annealing optimal

In this chapter, we show how *digitized*-QA can be made optimal, realizing the best possible solution allowed by quantum mechanics in the shortest time, without any prior knowledge on the location and properties of the spectral gap. We illustrate this on the simple benchmark problem of an unfrustrated antiferromagnetic Ising chain in a transverse field, where we provide an exact variational bound on the best accuracy achievable by *digitized*-QA. Our findings elucidate the intimate relation between *digitized*-QA, optimal Quantum Control, and recently proposed hybrid quantum-classical variational algorithms for quantum-state preparation and optimization.

4.1 Introduction

As discussed in Chapter 1, we tackle the issue of schedule optimization for *digitized*-QA from the perspective of hybrid quantum classical variational schemes.

We recall that schedule optimization for QA algorithms is believed to require, *in general*, spectral information on the problem at hand and this poses, clearly, a severe limitation [Ambainis, 2013; Cubitt *et al.*, 2015]. Hybrid quantum-classical variational algorithms, instead, are insensitive to critical points and spectral information. They are based on classical minimization and invoke quantum digital processors to prepare a variational *Ansatz* for the problem [Farhi *et al.*, 2014; Peruzzo *et al.*, 2014; Kokail *et al.*, 2019]. In the specific field of combinatorial optimization, this is accomplished by the Quantum Approximate Optimization Algorithm (QAOA) [Farhi *et al.*, 2014] that operates through a depth-P circuit of

digital unitary gates. In this framework, a problem is hard if it requires large-P (deep) quantum circuits to prepare a good *Ansatz*, or if the classical optimization landscape is complex and difficult to sample.

Although QA and QAOA appear as unrelated models of computation, they are both computationally universal [Aharonov *et al.*, 2008; Mizel *et al.*, 2007; Lloyd, 2018], suggesting that some connections might exist. Here we take a step forward in establishing this connection, by showing that one can construct an optimal QAOA solution which is *adiabatic*. Our contribution builds upon two recent interesting works. The first is the proposal for a fully *digitized*-QA (dQA) [Barends *et al.*, 2016]. As we discussed in Chapter 1 and Chapter 3, *digitized*-QA shares technical similarities with the QAOA quantum circuit and points towards a universal-gate approach to QA, with the bonus of the possibility of error-correction [Bravyi and Kitaev, 1998; Fowler *et al.*, 2012]. The second is the result of Yang *et al.* [2017], who showed that the digital nature of the QAOA *Ansatz* emerges naturally, when searching for an optimal protocol, from the "bang-bang" form predicted by the application of Pontryagin's principle [D'Alessandro, 2007; Brif *et al.*, 2010].

To demonstrate the construction of the optimal digitized-QA solution, we illustrate it for the benchmark problem of a quantum Ising chain, where a detailed size-scaling analysis is feasible. We start by proving a rigorous bound on the performance of any digital depth-P quantum circuit on such a problem. We then show that, among the apparently equivalent 2^{P} circuits that attain the best achievable performance, there is a special *regular optimal circuit* which can be constructed iteratively on P, and which can be interpreted as an *optimal digitized*-QA schedule. Such a schedule, obtained *without any spectral information*, is shown to be computationally optimal.

The rest of the Chapter is organized as follows. In Section 4.2 we recapitulate the methodological aspects of how to go from QA to QAOA, and recall the MaxCut problem in general. In Section 4.3, we introduce a technique, involving boundary conditions in a reduced chain, which allows us to prove the residual energy bound. In Sec. 4.4, with the aid of a Jordan-Wigner transformation, we cast the multivariate minimization into a much simpler geometric problem, and we explicitly show how to construct optimal *digitized*-QA adiabatic solutions, and compare it with other QA strategies. Finally, Sec. 4.5 contains a summary of our results, a discussion of relevant points, and an outlook on open questions.

4.2 Model and Methods

We start this section by reviewing, for the reader's, the connections between QA, *digitized*-QA, and QAOA, previously discussed in Chapter 1.

We consider the problem of preparing the ground state of a target Hamiltonian $\widehat{H}_{\mathrm{T}} = \widehat{H}_z$, which we assume, to keep the notation simple, to contain only $\hat{\sigma}^z$ terms. We supplement $\widehat{H}_z$ with the usual transverse field driving

$$\widehat{H}_x \;=\; -\sum_{j=1}^{N} \hat{\sigma}_j^x \,. \tag{4.1}$$

In the simplest setting [Albash and Lidar, 2018], we then write an interpolating QA Hamiltonian of the form (see Eq. (1.1)):

$$\widehat{H}(s) = s\,\widehat{H}_z + (1-s)\,\widehat{H}_x \,. \tag{4.2}$$

The parameter s is then varied in time, defining a schedule $s(t)$ interpolating from $s(0) = 0$ to $s(\tau) = 1$, where τ is the total annealing time. Given any $s(t)$, and starting from the ground state of $\widehat{H}_x$, $|\psi_0\rangle = 2^{-N/2}\left(|\uparrow\rangle + |\downarrow\rangle\right)^{\otimes N}$, the state of the system at time τ is given by the Schrödinger evolution $|\psi(\tau)\rangle = \widehat{U}_{\mathrm{QA}}(\tau,0)|\psi_0\rangle$ where $\widehat{U}_{\mathrm{QA}}(\tau,0)$ is the evolution operator, formally expressed as a time-ordered exponential, $\widehat{U}_{\mathrm{QA}}(\tau,0) = \mathrm{Texp}\left(-\frac{i}{\hbar}\int_0^\tau \mathrm{d}t'\,\widehat{H}(s(t'))\right)$. By approximating the schedule $s(t)$ with a step function attaining P values $s_{m=1,\cdots,\mathrm{P}} \in (0,1]$, and corresponding evolution times $\Delta t_{m=1,\cdots,\mathrm{P}}$ such that $\sum_{m=1}^{\mathrm{P}} \Delta t_m = \tau$ — see sketch in Fig. 4.1 —, and then taking a further Trotter splitting, the approximate evolution operator reads

$$\widehat{U}_{\mathrm{digit}} = \mathrm{e}^{-i\beta_{\mathrm{P}}\widehat{H}_x}\mathrm{e}^{-i\gamma_{\mathrm{P}}\widehat{H}_z}\cdots\mathrm{e}^{-i\beta_1\widehat{H}_x}\mathrm{e}^{-i\gamma_1\widehat{H}_z} \,, \tag{4.3}$$

where the parameters (for a lowest-order splitting $\gamma_m = s_m\Delta t_m/\hbar$ and $\beta_m = (1-s_m)\Delta t_m/\hbar$, see Eq. (1.23)) are such that:

$$\sum_{m=1}^{\mathrm{P}}(\beta_m+\gamma_m) = \frac{\tau}{\hbar} \,. \tag{4.4}$$

Eq. (4.3) naturally leads to the Quantum Approximate Optimization Algorithm (QAOA) introduced by Farhi *et al.* [2014]. Indeed, one can regard the quantum state

$$|\psi_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle = \mathrm{e}^{-i\beta_{\mathrm{P}}\widehat{H}_x}\mathrm{e}^{-i\gamma_{\mathrm{P}}\widehat{H}_z}\cdots\mathrm{e}^{-i\beta_1\widehat{H}_x}\mathrm{e}^{-i\gamma_1\widehat{H}_z}|\psi_0\rangle \,, \tag{4.5}$$

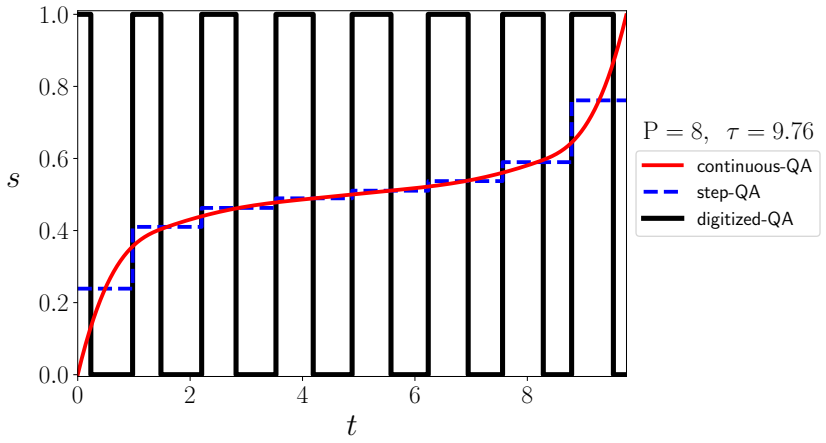

FIGURE 4.1: Optimal schedule for P = 8 (corresponding to $\tau = 9.76$). Lines: solid red, the continuous-time $s(t)$; dashed blue, the step-discretization; solid black, the bang-bang digitized-QA. The procedure by which this figure is generated is explained in Sec. 4.4.1.

as variationally dependent on the 2P parameters $(\boldsymbol{\gamma}, \boldsymbol{\beta})$. Using a quantum device that prepares $|\psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$ and performs repeated measurements in the computational basis, we then evaluate the expectation value of the target Hamiltonian

$$E_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle \psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) | \widehat{H}_z | \psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) \rangle \ , \tag{4.6}$$

and minimize it through a classical algorithm. The global variational minimum $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ determines a correspondingly optimal state $|\psi_{\mathrm{P}}(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)\rangle$, whose energy $E_{\mathrm{P}}^{\mathrm{opt}} = E_{\mathrm{P}}(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ is, by construction, a monotonically decreasing function of P.

We can quantify the degree to which a variational QAOA state $|\psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$ approximates the solution of the quantum problem with the rescaled residual energy(or relative error) [Santoro *et al.*, 2002]

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \frac{E_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) - E_{\mathrm{min}}}{E_{\mathrm{max}} - E_{\mathrm{min}}} \ , \tag{4.7}$$

where E_{min} and E_{max} are the minimum and maximum eigenvalues of $\widehat{H}_z$. $\epsilon_{\mathrm{P}}^{\mathrm{res}}$ is normalized [1] in such a way that $\epsilon_{\mathrm{P}}^{\mathrm{res}} \in [0, 1]$ and that $\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) = 0$ if and only if $|\psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$ is a ground state of $\widehat{H}_z$. This quantity is related to the approximation ratio r_{P} usually considered in the context of QAOA [Farhi *et al.*, 2014; Wang *et al.*, 2018; Pichler *et al.*, 2018] by the simple relation $r_{\mathrm{P}} = 1 - \epsilon_{\mathrm{P}}^{\mathrm{res}}$.

[1] Notice that the definition of $\epsilon_{\mathrm{P}}^{\mathrm{res}}$ differs from the one given in Chapter 2 (for continuous-time coherent-QA) by a slightly different normalization factor.

By evaluating r_{P} (or $\epsilon_{\mathrm{P}}^{\mathrm{res}}$) on the output of the QAOA we get the approximation ratio $r_{\mathrm{P}}^{*} = r_{\mathrm{P}}(\boldsymbol{\gamma}^{*}, \boldsymbol{\beta}^{*})$, a natural figure of merit for the performance of the algorithm [Farhi *et al.*, 2014]. Clearly, increasing the number of variables improves the QAOA's performance, hence $r_{\mathrm{P+1}}^{*} \geq r_{\mathrm{P}}^{*}$ or, equivalently, $\left(\epsilon_{\mathrm{P+1}}^{\mathrm{res}}\right)^{\mathrm{opt}} \leq (\epsilon_{\mathrm{P}}^{\mathrm{res}})^{\mathrm{opt}}$. Moreover, if the interpolating Hamiltonian $\widehat{H}(s)$ has a finite minimum gap in the interval $s \in [0, 1]$, an arbitrarily good solution can be obtained by Trotterization of an adiabatic evolution, so that $r_{\mathrm{P}}^{*} \to 1$ and $(\epsilon_{\mathrm{P}}^{\mathrm{res}})^{\mathrm{opt}} \to 0$ for $\mathrm{P} \to \infty$ [Farhi *et al.*, 2014].

Various studies of the performance of QAOA [Farhi *et al.*, 2014; Wang *et al.*, 2018; Pichler *et al.*, 2018; Crooks, 2018] have considered particular classes of Max-Cut graphs, which we introduced in Chapter 1. The simplest cases are connected k-regular graphs, where the connectivity (or degree) k, i.e., the number of edges originating from each node, is fixed and identical for all nodes.

A 2-regular graph is a closed ring, and the MaxCut problem — sometimes referred to as *ring of disagrees* [Farhi *et al.*, 2014] — is equivalent to finding the classical ground state of an anti-ferromagnetic Ising model on a chain, a computationally "easy" problem which, physically, shows no frustration (for N even). Generic k-regular graphs with $\mathrm{k} \geq 3$ are frustrated and solving the associated Max-Cut problem is hard. For 3-regular graphs, Farhi *et al.* [2014] has shown that the *worst case* value of the approximation ratio r_{P} is bound, for $\mathrm{P} = 1$, to be $r_{\mathrm{P=1}}^{*} \geq 0.6924$.

To illustrate our results, we will now consider the MaxCut on a 2-regular graph, *alias* the anti-ferromagneticIsing chain:

$$\widehat{H}_z \;=\; \sum_{j=1}^{N} (\hat{\sigma}_j^z \hat{\sigma}_{j+1}^z - 1) \tag{4.8}$$

where N (which we assume to be even) is the number of sites in the chain, and PBC are assumed.

For this problem, we will now show that:

1 One can prove a lower bound for the residual energy $\epsilon_{\mathrm{P}}^{\mathrm{res}}$ or, equivalently an upper bound for r_{P}:

$$\epsilon_{\mathrm{P}}^{\mathrm{res}} \geq \begin{cases} \dfrac{1}{2\mathrm{P}+2} & \text{for } 2\mathrm{P} < N \\ 0 & \text{for } 2\mathrm{P} \geq N \end{cases} . \tag{4.9}$$

2 By mapping the problem to free fermions, through Jordan-Wigner, we verify numerically that the bound in Eq. (4.9) is saturated (at least for $\mathrm{P} \leq 128$). This extends the results of Wang *et al.* [2018], limited to $\mathrm{P} \leq 10$. For $2\mathrm{P} \geq N$ we explicitly exhibit an optimal schedule attaining $\epsilon_{\mathrm{P}}^{\mathrm{res}} = 0$, see (4.31).

3 Among the many degenerate "structureless" optimal solutions, we are able to construct a "regular" representative $(\boldsymbol{\gamma}^{\mathrm{reg}}, \boldsymbol{\beta}^{\mathrm{reg}})$. We show that it can be associated to a smooth QA schedule $s(t)$, thus realizing an *optimal digitized*-QA dynamics.

4.3 Variational bound on QAOA performance

We show here that the upper bound in Eq. (4.9) applies to the 2-regular graph, i.e., the antiferromagnetic Ising chain, when $(2\mathrm{P}+2) \leq N$.

The translational invariance of $\widehat{H}_z$, $\widehat{H}_z$ and of the initial state $|\psi_0\rangle$ implies that we can write the residual energy in Eq. (4.7) as:

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle \psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) | \frac{\hat{\sigma}_{j_s}^z \hat{\sigma}_{j_s+1}^z + 1}{2} | \psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) \rangle \ , \tag{4.10}$$

where j_s is any site of the chain, for instance the central site $j_s = N/2$. [2] As demonstrated by Farhi *et al.* [2014], the application of the digital unitary operator $\widehat{U}_{\mathrm{digit}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \widehat{U}_{\mathrm{P}} \cdots \widehat{U}_1$ — where $\widehat{U}_m = \widehat{U}(\gamma_m, \beta_m)$ — to link operator $\hat{\sigma}_{j_s}^z \hat{\sigma}_{j_s+1}^z$

$$\widehat{U}_1^\dagger \cdots \widehat{U}_{\mathrm{P}}^\dagger \, \hat{\sigma}_{j_s}^z \hat{\sigma}_{j_s+1}^z \, \widehat{U}_{\mathrm{P}} \cdots \widehat{U}_1 \tag{4.11}$$

involves only spins which have a distance at most P from the link (j_s, j_s+1). Such an operator spreading is sketched in Fig. 4.2, where a certain similarity with the light-cone idea emerges [Yuezhen Niu *et al.*, 2019]. Considering for instance the central link $j_s = \frac{N}{2}$, if $(2\mathrm{P}+2) < N$, this leads to a reduced spin chain with $N_{\mathrm{R}} = (2\mathrm{P}+2)$ sites, $j = j_s - \mathrm{P}, \cdots, j_s + \mathrm{P} + 1$, while if $(2\mathrm{P}+2) \geq N$, all sites are involved, $N_{\mathrm{R}} = N$. Without loss of generality, we can always re-number the sites belonging to the reduced spin chains from 1 to N_{R} and assume that the link we are "measuring" is the central one, $j_s = N_{\mathrm{R}}/2$.

Notice that for $(2\mathrm{P}+2) = N_{\mathrm{R}} \leq N$, in the spirit of the spin reduction explained above, any boundary term is *absent* in the reduced chain Hamiltonian. Hence, we

[2] We restrict ourself to even N so that the chain is unfrustrated.

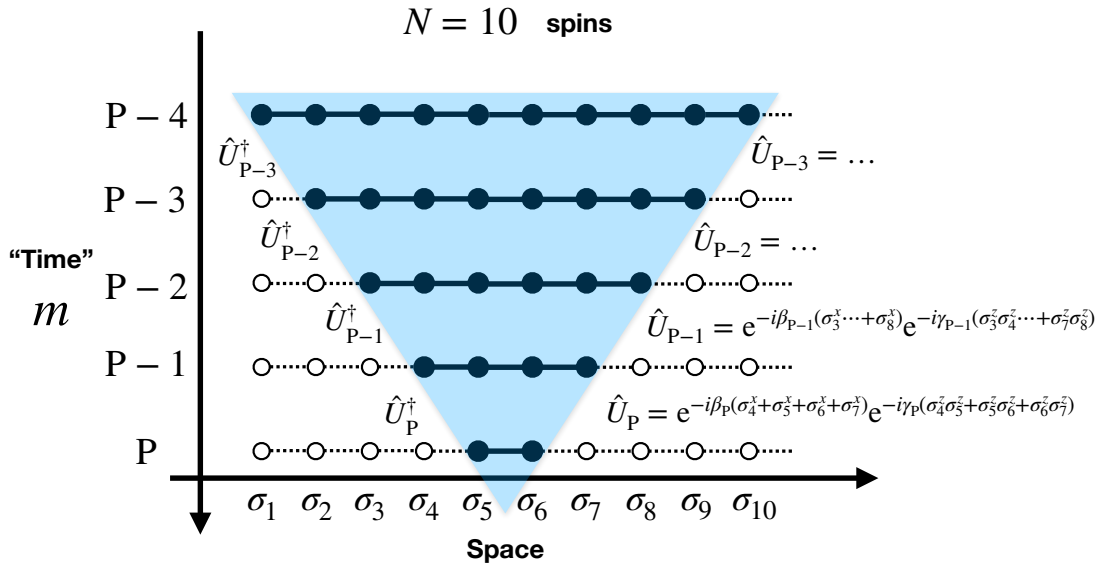

FIGURE 4.2: A sketch of how the successive application of digital evolution results in an operator spreading that justifies the use of a reduced spin chain. Notice the systematic absence of boundary terms in the reduced chain Hamiltonian.

are free to add an *arbitrary* boundary link term $J_b\,\hat{\sigma}^z_{N_\mathrm{R}}\hat{\sigma}^z_1$. We restrict our choice here to $J_b = \pm 1$. For $J_b = +1$ the reduced spin chain has PBC, hence recovering full translational invariance. For $J_b = -1$ we have anti-periodic boundary conditions (ABC), but an effective translation operator involving a spin-flip can still be introduced (see App. C). Our claim now is that we can work with a reduced spin chain Hamiltonian of the form:

$$\hat{\mathcal{H}}^{(\pm)}_z = \sum_{j=1}^{N_\mathrm{R}-1}\left(\hat{\sigma}^z_j\hat{\sigma}^z_{j+1} - 1\right) + \left(\pm\,\hat{\sigma}^z_{N_\mathrm{R}}\hat{\sigma}^z_1 - 1\right)\,, \tag{4.12}$$

while keeping the transverse term unmodified

$$\hat{\mathcal{H}}_x = -\sum_{j=1}^{N_\mathrm{R}}\hat{\sigma}^x_j\,, \tag{4.13}$$

and show that this would not modify the expectation value we need, i.e.,

$$\epsilon^{\mathrm{res}}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta}) = \langle\widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta})|\frac{\hat{\sigma}^z_{j_s}\hat{\sigma}^z_{j_s+1}+1}{2}|\widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle \tag{4.14}$$

where

$$|\widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle = \prod_{m=1}^{\leftarrow\mathrm{P}}\mathrm{e}^{-i\beta_m\hat{\mathcal{H}}_x}\mathrm{e}^{-i\gamma_m\hat{\mathcal{H}}^{(\pm)}_z}\,|\widetilde{\psi}_0\rangle\,, \tag{4.15}$$

with $|\widetilde{\psi}_0\rangle = |+\rangle^{\otimes N_{\rm R}}$. Such extra freedom in the boundary conditions will be used shortly to derive a bound for the residual energy. We notice that when $2\mathrm{P}+2 > N$ the reduced spin chain coincides with the full chain, $N_{\rm R} = N$, and there is no freedom whatsoever: we must use PBC.

Returning to the case $2\mathrm{P}+2 = N_{\rm R} \leq N$, the next crucial step is to show that:

$$\langle\widetilde{\psi}_{\rm P}|\left(\hat{\sigma}^z_{j_s}\hat{\sigma}^z_{j_s+1} - 1\right)|\widetilde{\psi}_{\rm P}\rangle \;=\; \frac{1}{N_{\rm R}}\langle\widetilde{\psi}_{\rm P}|\hat{\mathcal{H}}^{(\pm)}_z|\widetilde{\psi}_{\rm P}\rangle\;. \tag{4.16}$$

For the PBC case, this is a trivial consequence of translational invariance. For the ABC case, one needs to show that a modified translation operator can be introduced, by incorporating a spin-flip at site 1, which does the job: this is shown explicitly in App. C. Using Eqs. (4.16) and (4.14) we immediately conclude that

$$\begin{aligned}\epsilon^{\rm res}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta}) &= \langle\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})|\left(\frac{\hat{\mathcal{H}}^{(\pm)}_z}{2(2\mathrm{P}+2)}+1\right)|\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle \\ &\geq \left(\frac{E^{(\pm)}_{\rm gs}}{2(2\mathrm{P}+2)}+1\right)\;,\end{aligned} \tag{4.17}$$

where $E^{(\pm)}_{\rm gs}$ is the ground state energy of $\hat{\mathcal{H}}^{(\pm)}_z$ and the inequality follows from the standard variational principle. Now observe that $E^{(+)}_{\rm gs} = -2N_{\rm R} = -2(2\mathrm{P}+2)$. Hence, if PBC are used in the reduced spin chain, the bound is a trivial useless inequality $\epsilon^{\rm res}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta}) \geq 0$. The inequality, however, becomes non-trivial if ABC are used, since $E^{(-)}_{\rm gs} = -4\mathrm{P}-2$ due to the frustrating boundary term $J_b = -1$:

$$\epsilon^{\rm res}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta}) \geq \left(\frac{E^{(-)}_{\rm gs}}{2(2\mathrm{P}+2)}+1\right) = \frac{1}{2\mathrm{P}+2}\;, \tag{4.18}$$

This establishes the promised bounds in Eq. (4.9).

As a check, observe that the computation of the optimal $r^*_{\mathrm{P}=1}$ carried out in Wang *et al.* [2018] (sse their Eq. (16)) and valid for triangle-free k-regular graphs, when translated into our notation would imply that:

$$\epsilon^{\rm res}_{\mathrm{P}=1} \geq \frac{1}{2} - \frac{\left(1-\frac{1}{\mathrm{k}}\right)^{\frac{\mathrm{k}-1}{2}}}{2\sqrt{\mathrm{k}}}\;. \tag{4.19}$$

This shows that the optimal result in Eq. (4.19), when specialized to $\mathrm{k} = 2$, coincides with Eq. (4.18). In the next section, see Fig. 4.4 and accompanying discussion, we will explicitly demonstrate numerically the tightness of the bound

in Eq. (4.18) also for larger P.

The derivation of the bound in Eq. (4.9) given here relies on the locality and translational invariance of the problem. As we will see in Chapter 5 it can be generalized to regular graphs of higher order and more general ground state preparation problems.

4.4 Numerical results and tightness of the bound

For the 2-regular graph case, i.e., the antiferromagnetic Ising chain case, a lot more can be said. First of all, as discussed by Wang *et al.* [2018], one can take advantage of the Jordan-Wigner transformation [Jordan and Wigner, 1928] to map the problem into a free-fermion one. In particular, one can show that the system is equivalent to a set of independent two-level systems.

Let us recall, for the reader's convenience, that if $2\mathrm{P}+2 > N$, then we are dealing with a standard Ising chains of length $N_{\mathrm{R}} = N$ with the original (PBC) boundary conditions; otherwise, if $2\mathrm{P}+2 = N_{\mathrm{R}} \leq N$, the reduced spin chain discussed in the previous section can be taken to have an *arbitrary* boundary term. As discussed previously, using ABC leads to the bound we have derived. Hence, $N_{\mathrm{R}} = \min(N, 2\mathrm{P}+2)$ is the effective chain length, and we will set the boundary condition appropriately.

Using a Jordan-Wigner transformation (see App. A), the relevant Hamiltonians can be expressed as a sum of independent two-level systems labeled by a wave-vector k whose values depend on the boundary conditions used. In particular, we define $\mathcal{K}_{\mathrm{PBC}} = \{\frac{\pi}{N_{\mathrm{R}}}, \frac{3\pi}{N_{\mathrm{R}}} \cdots, \frac{(N_{\mathrm{R}}-1)\pi}{N_{\mathrm{R}}}\}$ and $\mathcal{K}_{\mathrm{ABC}} = \{\frac{2\pi}{N_{\mathrm{R}}}, \frac{4\pi}{N_{\mathrm{R}}}, \cdots, \frac{(N_{\mathrm{R}}-2)\pi}{N_{\mathrm{R}}}\}$ to be the set of k-vectors associated to PBC and ABC, respectively, for the spin chain. The final result, see App. A, is that the residual energy can also be decomposed, for $2\mathrm{P} < N$ as:

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\boldsymbol{\beta}) \overset{2\mathrm{P}\leq N}{=} \frac{1}{2\mathrm{P}+2} + \frac{1}{2\mathrm{P}+2}\sum_{k}^{\mathcal{K}_{\mathrm{ABC}}} \epsilon_k(\boldsymbol{\gamma},\boldsymbol{\beta}) \,, \tag{4.20}$$

while for $2\mathrm{P} \geq N$ we get:

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\boldsymbol{\beta}) \overset{2\mathrm{P}\geq N}{=} \frac{1}{N}\sum_{k}^{\mathcal{K}_{\mathrm{PBC}}} \epsilon_k(\boldsymbol{\gamma},\boldsymbol{\beta}) \,. \tag{4.21}$$

Here, the non-negative function $\epsilon_k(\boldsymbol{\gamma},\boldsymbol{\beta}) \geq 0$ is

$$\epsilon_k(\boldsymbol{\gamma},\boldsymbol{\beta}) = 1 - \hat{\boldsymbol{b}}_k^T \left(\prod_{m=1}^{\leftarrow \mathrm{P}} \mathcal{R}_{\hat{\boldsymbol{z}}}(4\beta_m)\mathcal{R}_{\boldsymbol{v}_k}(4\gamma_m) \right) \hat{\boldsymbol{z}} \,, \tag{4.22}$$

and represents the k-contribution to the residual energy, expressed in terms 3×3 rotation matrices $\mathcal{R}$ around unit vectors $\hat{\boldsymbol{z}} = (0,0,1)^T$ and $\hat{\boldsymbol{b}}_k = (-\sin k, 0, \cos k)^T$ by rotation angles $4\beta_m$ and $4\gamma_m$, respectively.

Eqs. (4.20), (4.21) and (4.22) are our starting points to discuss the properties of the QAOA landscape. The first observation is that the landscape has periodicity of $\pi/2$ in each variable γ_m and β_m [Farhi *et al.*, 2014]. Without loss of generality we can assume $\gamma_m, \beta_m \in [0, \frac{\pi}{2}]$.

A second observation emerges from the inspection of Eq. (4.20). For $2\mathrm{P} < N$ we use ABC and therefore $N_{\mathrm{R}} = 2\mathrm{P} + 2$, which implies that $\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\boldsymbol{\beta})$ is totally independent of N. This N-independence is, in retrospective, a general consequence of the spin reduction behind QAOA for translational invariant models [Farhi *et al.*, 2014], valid well beyond the Jordan-Wigner framework used to derive Eq. (4.20). Moreover, the optimal residual energy $\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*)$ saturates the bound in Eq. (4.9), hence $\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma}^*,\boldsymbol{\beta}^*) = \frac{1}{2\mathrm{P}+2}$, provided we are able to make the contribution from $\sum_k \epsilon_k$ to vanish (which we can, as discussed below).

The simple transformation properties of the system translate into corresponding properties for the QAOA landscape. In particular, one can show [Wang *et al.*, 2018; Mbeng *et al.*, 2019] that:

$$\begin{aligned} \epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\boldsymbol{\beta}) &= \epsilon_{\mathrm{P}}^{\mathrm{res}}(-\boldsymbol{\gamma},-\boldsymbol{\beta}) \\ \epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\boldsymbol{\beta}) &= \epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\beta}',\boldsymbol{\gamma}') \,, \end{aligned} \tag{4.23}$$

where we have defined the vectors $\boldsymbol{\beta}', \boldsymbol{\gamma}'$ as $\beta'_m = \beta_{\mathrm{P}-m-1}$ and $\gamma'_m = \gamma_{\mathrm{P}-m-1}$. In Wang *et al.* [2018] it was shown that the optimal values for the parameter lie in the sub-manifold $\boldsymbol{\beta} = \boldsymbol{\gamma}'$ for $\mathrm{P} \leq 10$. We have confirmed this result — which applies to the case $2\mathrm{P} < N$ — for $\mathrm{P} \leq 128$.

The function $\epsilon_k(\boldsymbol{\gamma},\boldsymbol{\beta})$ has a simple geometrical interpretation: it contains the scalar product of $\hat{\boldsymbol{b}}_k$ with the vector $\left(\prod_{m=1}^{\leftarrow \mathrm{P}} \mathcal{R}_{\hat{\boldsymbol{z}}}(4\beta_m)\mathcal{R}_{\boldsymbol{v}_k}(4\gamma_m)\right) \hat{\boldsymbol{z}}$ obtained by applying 2P successive rotations to $\hat{\boldsymbol{z}}$. Therefore ϵ_k assumes its minimum value 0 when

$$\left(\prod_{m=1}^{\leftarrow \mathrm{P}} \mathcal{R}_{\hat{\boldsymbol{z}}}(4\beta_m)\mathcal{R}_{\hat{\boldsymbol{b}}_k}(4\gamma_m) \right) \hat{\boldsymbol{z}} = \hat{\boldsymbol{b}}_k \,. \tag{4.24}$$

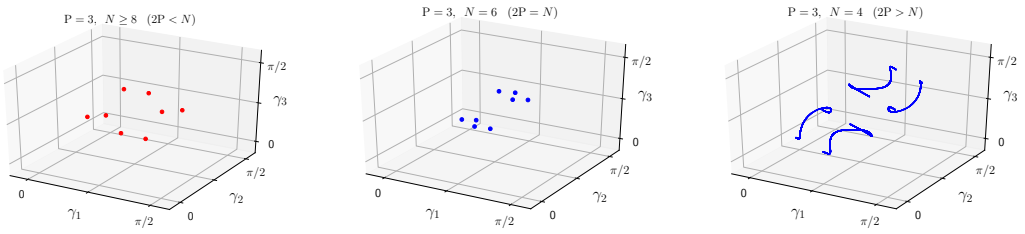

FIGURE 4.3: Visualization of the optimal solutions of QAOA for P = 3 in the symmetric manifold $\boldsymbol{\beta} = \boldsymbol{\gamma}'$. Blue circles correspond to values $\epsilon^{\rm res}_{\rm P=3} = 0$ while red ones corresponds to the values $\epsilon^{\rm res}_{\rm P=3} = 1/(2{\rm P}+2) = 1/8$. In (a), for $2{\rm P} < N$ ($N \geq 8$), the optimization problem has a finite set of solutions corresponding to strictly positive values of $\epsilon^{\rm res} = 1/8$. In (b), for $2{\rm P} = N$ ($N = 6$), there is a finite set of optimal solutions, now having $\epsilon^{\rm res} = 0$. In (c), for $2{\rm P} > N$ (here $N = 4$), the manifold of solutions attaining $\epsilon^{\rm res} = 0$ has dimension $2{\rm P} - N = 2$: the curves shown are obtained by intersecting the solution surface with the symmetric manifold $\boldsymbol{\beta} = \boldsymbol{\gamma}'$.

These can be regarded as a set of *constraints*, whose number depends on the number of k-vectors involved. The minimum residual energy is obtained by simultaneously minimizing, whenever possible, all the addends of the k-sum in Eqs. (4.20)-(4.21). By counting the number of free variational parameters and the number of constraint equations, one can get a picture of the QAOA landscape. Figure 4.3 illustrates the role of these counting arguments for P = 3. Here we observe that:

For 2P < N: the number of constraints $2|\mathcal{K}_{\rm ABC}| = 2{\rm P}$ — corresponding to P equations (4.24) for 3-dimensional unit vectors — and of variables 2P are equal. Therefore the equations have a finite set of discrete solutions. When all equations are satisfied we get, see (4.20), the optimal $\epsilon^{\rm res}(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) = (2{\rm P}+2)^{-1}$.

For 2P = N: again the number of constraints $2|\mathcal{K}_{\rm PBC}| = N$ and of variables 2P are equal. Therefore the equations still have discrete solutions. When all equations are satisfied, the residual energy, see (4.21), is $\epsilon^{\rm res}(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) = 0$.

For 2P > N: the number of constraints $2|\mathcal{K}_{\rm PBC}| = N$ is smaller than the number of variables 2P. The equations therefore have a continuous set of solutions that define a manifold of dimension $2{\rm P} - N$.

Figure 4.4 illustrates the minimum residual energy obtained for different values of N as a function of P. The numerical data are obtained by looking for optimal solutions via a numerical minimization of the residual energy, Eq. (4.20), using Eq. (4.22) to compute the terms ϵ_k. Specifically we implement the function $\epsilon^{\rm res}_{\rm P}(\boldsymbol{\gamma}, \boldsymbol{\beta})$ with `PyTorch` that provides built-in auto-differentiation routines [Paszke

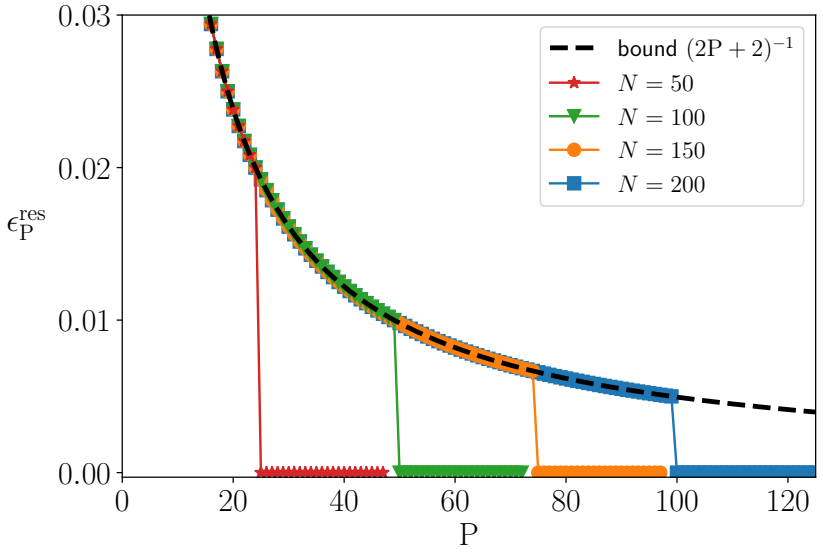

FIGURE 4.4: Optimal residual energies $\epsilon_{\mathrm{P}}^{\mathrm{res}}$ vs the number of steps P for the quantum Ising chain, at various system sizes $N = 50, 100, 150, 200$. The symbols represent the data obtained by numerical minimization and the dashed black line represents the theoretical bound in (4.9). The bound is saturated for $2\mathrm{P} < N$. For $2\mathrm{P} \geq N$, the residual energy drops to zero.

et al., 2017]. We then minimize $\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma}, \boldsymbol{\beta})$ with the Broyden-Fletcher-Goldfard-Shanno (BFGS) optimization algorithm [Nocedal and Wright, 2006], using back-propagation to compute the required gradients. The algorithm is halted when the residual energy is sufficiently close — specifically, within 10^{-7} — to our theoretical lower bound $\epsilon_{\mathrm{P}}^{\mathrm{res}} = \frac{1}{2\mathrm{P}+2}$.

The global minima returned by the BFGS routine depend on the arbitrary choice of the initial guess for $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$. In particular, there are 2^{P} *degenerate minima* all sharing the same $\epsilon_{\mathrm{P}}^{\mathrm{res}} = \frac{1}{2\mathrm{P}+2}$ for $N > 2\mathrm{P}$. Notice that $\epsilon_{\mathrm{P}}^{\mathrm{res}}$ drops to 0 when $2\mathrm{P} \geq N$, as predicted by the parameter counting argument presented above.

We will now show that, among all these degenerate solutions, one can single-out a rather special *regular solution* which is closely related to the problem of an optimal QA [Rezakhani *et al.*, 2009].

4.4.1 Optimal schedules for $2\mathrm{P} < N$ and digitized-QA

As discussed previously, for $2\mathrm{P} < N$ the QAOA landscape is independent on the system size, and one is effectively considering an infinite chain $N \to \infty$. As we already mentioned there are various equivalent optimal choices for the γ_m and β_m, most of which lack any structure or pattern. In this section we exhibit a *regular* schedule that shows a well defined continuous limit when $\mathrm{P} \to \infty$.

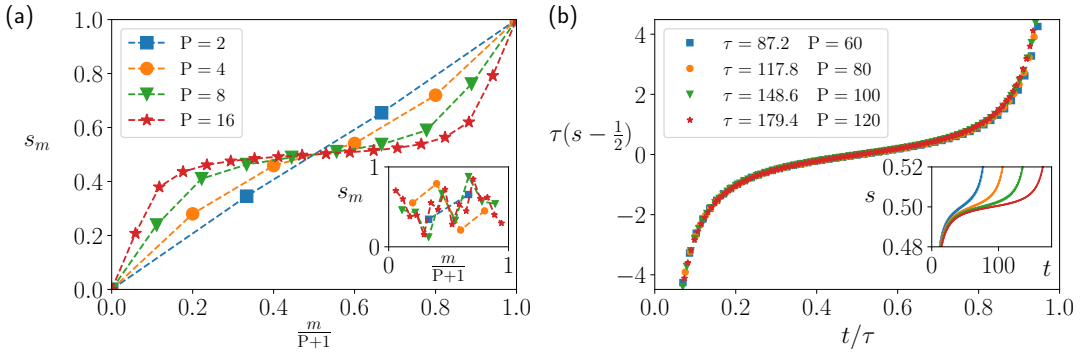

FIGURE 4.5: (a) Construction of the "regular" solution for increasing P. We plot here the parameter $s_m = \frac{\gamma_m}{\gamma_m+\beta_m}$. The regular solution for P Trotter steps is obtained using as an initial guess for the local search algorithm the solution obtained for a smaller value of P (e.g. P/2). This procedure produces a smooth schedule. The inset shows a set of generic optimal solutions obtained when initializing the QAOA angles randomly. This procedure apparently produces irregular patternless solutions. (b) Control protocol $s(t)$ induced by the QAOA algorithm for several values of the total time τ (or equivalently the number of Trotter steps P). The protocols have been scaled according to Eq. (4.26). The inset shows the unscaled protocols.

To specifically target the regular solution, we proceed iteratively in P [Pichler *et al.*, 2018]. The optimal solution at level P is obtained by using as an initial guess for $\boldsymbol{\gamma}, \boldsymbol{\beta}$ the regular solution previously obtained at level $\mathrm{P}' < \mathrm{P}$. For instance, starting from $\mathrm{P} = 2$ we can then iteratively obtain regular solutions for larger values of P, as illustrated in Fig. 4.5. For $\mathrm{P} = 2$, the solution, when represented in terms of the effective s parameter obtained by inverting Eq. (1.23)

$$s_m = \frac{\gamma_m}{\gamma_m + \beta_m} \tag{4.25}$$

coincides to a good approximation with the expected "linear interpolation" between $s = 0$ and $s = 1$, a standard choice in ordinary continuous-time QA, which is here used as initial starting point in the BFGS search for the minimum. We next consider $\mathrm{P} = 4$ and start the BFGS minimization search from the interpolation of the $\mathrm{P} = 2$ values. The minimum found deviates now from the "linear interpolation". Proceeding further, we get the solutions shown in the main plot of Fig. 4.5(a), whose inset, by contrast, illustrates the "irregular" values of s_m that one would obtain, for the same P, by starting the search from a random initial point. Summarizing, among the vast majority of irregular solutions, one can single out, through an appropriate iterative search scheme, a *regular* solution whose parameters s_m appear to have a well recognizable "structure", which is further analyzed in Fig. 4.5(b). The inset of Fig. 4.5(b) shows that by increasing P, hence

the total τ given by Eq. (4.4), the schedule s_m, when expressed in terms of the corresponding time $t_m = \sum_{m'=1}^{m}(\gamma_{m'} + \beta_{m'})/\hbar$, appears to become flatter and flatter close to the quantum critical point $s = 1/2$. Remarkably, the whole regular solution shows a simple scaling of the form

$$s_\tau(t) = \frac{1}{2} + \frac{1}{\tau^\alpha} f\left(\frac{t}{\tau}\right) \tag{4.26}$$

with $\alpha = 1$, as illustrated by the data collapse in the main plot of Fig. 4.5(b). Such a shape of $s(t)$ is clearly reminiscent of the adiabatic protocols described in the context of a continuous-time QA in Roland and Cerf [2002] or Barankov and Polkovnikov [2008]. We will further comment on such an issue in Sec. 4.4.3.

We now explore the connection with QA, or more properly to a form of *digitized*-QA [Barends *et al.*, 2016]. The optimal parameters of the regular solution $(\boldsymbol{\gamma}^{\rm reg}, \boldsymbol{\beta}^{\rm reg})$ define a candidate *digitized*-QA schedule, from which, by inverting Eq. (1.23), one can construct an associated step-QA and a continuous-time QA protocol, as illustrated in Fig. 4.1 for $\mathrm{P} = 8$.

Our iterative construction targets an optimal solution $\boldsymbol{\gamma}^{\rm reg}, \boldsymbol{\beta}^{\rm reg}$ that varies weakly from P to $\mathrm{P}+1$. However, to explore the connection with a *digitized*-QA, we also need to verify and quantify the adiabaticity of the dynamics defined by $\boldsymbol{\gamma}^{\rm reg}, \boldsymbol{\beta}^{\rm reg}$.

During the preparation of the variational state given in Eq. (4.5), the system undergoes a unitary discrete time evolution. The intermediate state $|\psi_{m+1,\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$ after m steps satisfies the following discrete version of Schrödinger's equation

$$|\psi_0(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = |\psi_0\rangle \tag{4.27}$$

$$|\psi_{m+1}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = \widehat{U}_m |\psi_m(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle \tag{4.28}$$

where we recall that the effective discrete time evolution operator is defined to be $\widehat{U}_m = \widehat{U}(\gamma_m, \beta_m) = \mathrm{e}^{-i\beta_m \widehat{H}_x} \mathrm{e}^{-i\gamma_m \widehat{H}_{\rm T}}$. We can always find an orthonormal basis $|\theta_m\rangle$ that diagonalizes $\widehat{U}_m$:

$$\widehat{U}_m |\theta_m\rangle = \mathrm{e}^{-i\theta_m} |\theta_m\rangle \tag{4.29}$$

We say that $(\boldsymbol{\gamma}, \boldsymbol{\beta})$ defines an adiabatic dynamics if the state $|\psi_m(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$ closely *follows an eigenstate* $|\overline{\theta}_m\rangle$ of $\widehat{U}_m$. This is a natural extension [Dranov *et al.*, 1998] of the concept of adiabadicity in continuous-time dynamics. For instance, one can show that a *digitized*-QA schedule obtained by discretizing a continuous-QA (as

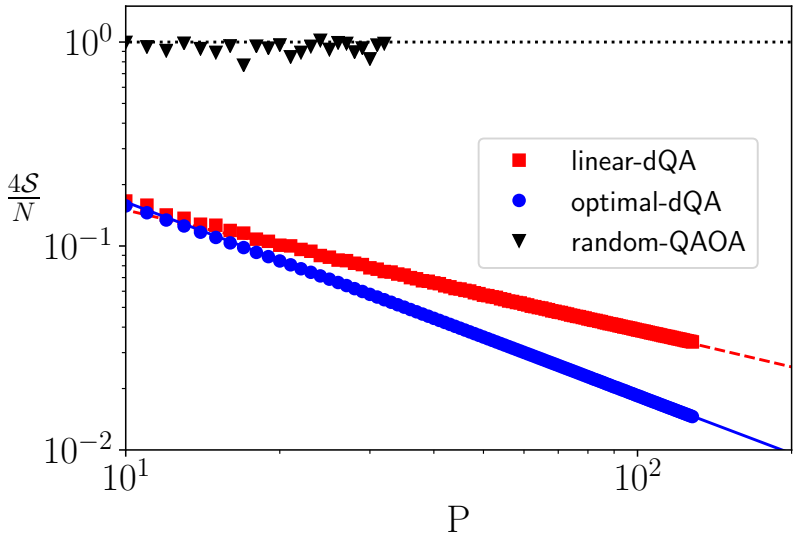

FIGURE 4.6: Average Shannon entropy $\mathcal{S}_{\boldsymbol{\gamma},\boldsymbol{\beta}}$ defined in Eq. (4.30), normalized to $N/4$, for various schedules.

the one studied in Chapter 3) is as adiabatic as its continuous counterpart when $\mathrm{P} \to \infty$. Moreover, this definition is further justified by the statements in Dranov *et al.* [1998], where a discrete version of the adiabatic theorem is given.

Let $p_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\theta_m) = |\langle\theta_m|\psi_m(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle|^2$ be the probability of finding the system in a given eigenstate $|\theta_m\rangle$. The definition of adiabatic dynamics given above, suggests to quantify the degree of adiabaticy by measuring how close the distribution $p_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\theta_m)$ is to a degenerate one (i.e., a Kröneker-delta in θ_m). The adiabaticity of the discrete dynamics with P steps, can then be quantified with the average Shannon entropy $\mathcal{S}_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\mathrm{P})$ of the distribution $p_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\theta_m)$:

$$\mathcal{S}_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\mathrm{P}) = -\frac{1}{\mathrm{P}}\sum_{m=1}^{\mathrm{P}}\sum_{\theta_m} p_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\theta_m)\log[p_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\theta_m)] \tag{4.30}$$

For an adiabatic dynamics $\mathcal{S}_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\mathrm{P}) \to 0$ as $\mathrm{P} \to \infty$, otherwise it should remain finite. In Fig. 4.6 we show such Shannon entropy for three different schedules on an Ising chain with $N = 1024$ sites. We first take a look at $\mathcal{S}_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\mathrm{P})$ for a linear *digitized*-QA schedule ($\Delta t_m = 1$) which is represented by red squares. Through the adiabatic theorem we know that the linear-digital schedule is adiabatic for $\mathrm{P} \to \infty$ and $\mathcal{S}_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\mathrm{P})$, decaying to zero when P is increased, correctly signals the emerging adiabaticity of the schedule. We then consider a generic optimal solution found by the QAOA algorithm staring form a random initialization (black triangles). We find that $4\mathcal{S}_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\mathrm{P})/N \approx 1$ independently of P, signalling a non-adiabatic dynamics. Finally, the blue circles were obtained from the regular QAOA

solution $\boldsymbol{\gamma}^{\rm reg}, \boldsymbol{\beta}^{\rm reg}$. The fact that $\mathcal{S}_{\boldsymbol{\gamma}^{\rm reg},\boldsymbol{\beta}^{\rm reg}}(\mathrm{P})$ vanishes as $\mathrm{P} \to \infty$ immediately conveys the message that the regular QAOA solution defines an adiabatic schedule. Moreover $\mathcal{S}_{\boldsymbol{\gamma},\boldsymbol{\beta}}(\mathrm{P})$ allows us to make quantitative statements: In particular, the regular QAOA solution is evidently *more* adiabatic than the linear *digitized*-QA schedule. We conclude that such an optimal solution can be interpreted as an improved adiabatic *digitized*-QA schedule. In App. D, we discuss how a suitable effective Hamiltonian can be introduced for the *digitized*-QA.

4.4.2 Optimal schedules for $2\mathrm{P} \geq N$ and Quantum Control

For $2\mathrm{P} \geq N$ it is always possible to prepare the Ising $\widehat{H}_{\rm T}$ ground state with a QAOA *Ansatz*: the system is controllable. This can be done by explicitly showing that the specific schedule

$$\beta_m = \gamma_{\mathrm{P}+1-m} = \begin{cases} \pi/8 & \text{if } m = \lceil \frac{\mathrm{P}+1}{2} \rceil \\ \pi/4 & \text{otherwise} \end{cases} \tag{4.31}$$

realizes exactly $\epsilon_{\rm P}^{\rm res} = 0$. The rationale behind such a remarkably simple expression is that most of the rotations $\mathcal{R}_{\hat{\boldsymbol{z}}}(4\beta_m)\mathcal{R}_{\boldsymbol{v}_k}(4\gamma_m)$ involved in Eq. (4.22) are rotations by π and their combined effect leaves the vector $\hat{\boldsymbol{b}}_k$ in the same plane as $\hat{\boldsymbol{z}}$ while shifting the angle between them by $2k$. The discrete nature of the k-vectors involved guarantees that one effectively rotates, by using the angles in Eq. (4.31), each $\hat{\boldsymbol{b}}_k$ onto $\hat{\boldsymbol{z}}$, as a detailed construction (not given here) shows. More generally, however, since the problem is now underdetermined (the number of variational parameters is larger than the number of constraints), one can construct a *continuum* of optimal solutions. In particular, using the same iterative strategy described in the previous section, we can single-out a regular solution attaining $\epsilon_{\rm P}^{\rm res}(\boldsymbol{\gamma}^{\rm reg}, \boldsymbol{\beta}^{\rm reg}) = 0$. Figure 4.7(a) shows the construction of such a regular schedule for $2\mathrm{P} = N$. Concerning a collapse of the data, we verified that the *Ansatz* in Eq. (4.26) still works, but now with a modified exponent $\alpha \approx 1.75$. On the practical side, observe that the schedule becomes flatter and flatter across the quantum critical point ($s = 1/2$).

4.4.3 Comparison with other QA strategies

One might ask how the optimal regular digitized-QA solution compares with other standard QA approaches for the ordered Ising chain problem. Specifically, one

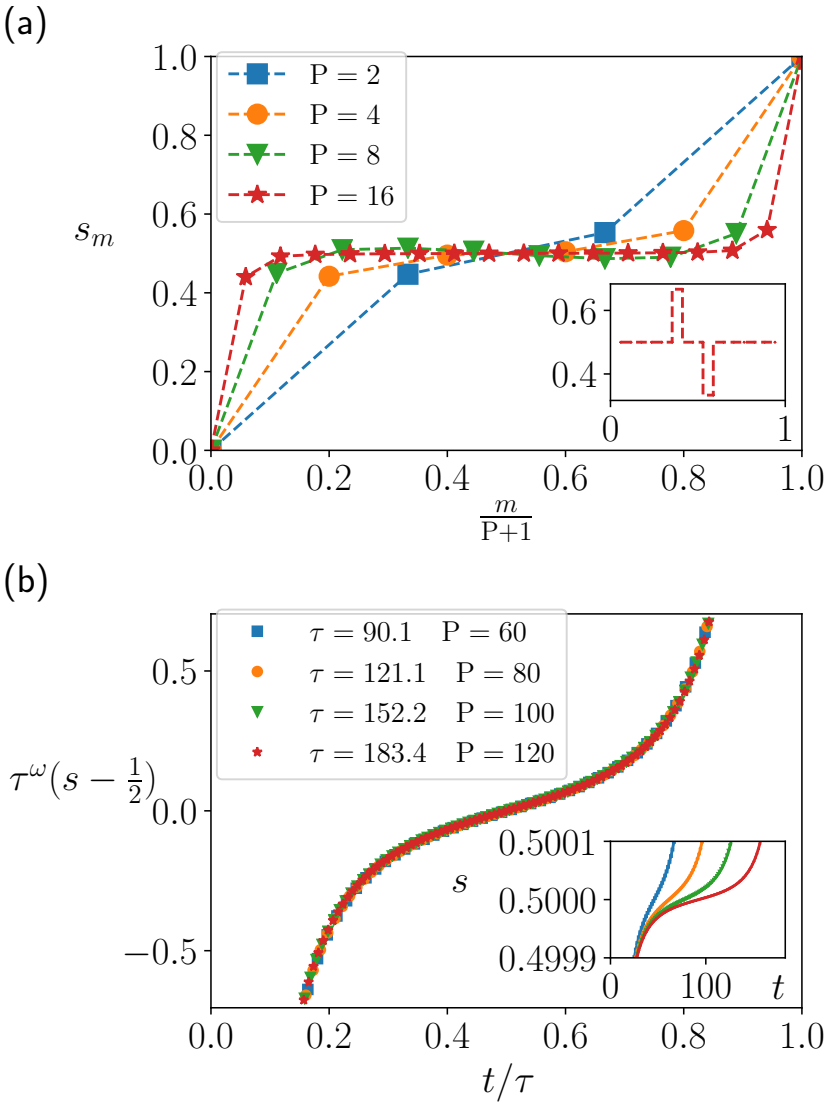

FIGURE 4.7: (a) Construction of the "regular" solution for increasing P when $2\mathrm{P} \geq N$. We plot here the parameter $s_m = \frac{\gamma_m}{\gamma_m + \beta_m}$. The regular solution for P Trotter steps is obtained using as initial guess for the local search algorithm the solution obtained for a smaller value of P (e.g. P/2). This procedure produces a smooth schedule. The inset shows the optimal solution in Eq. (4.31). (b) Control protocol $s(t)$ induced by the QAOA algorithm for several values of the total time τ (or, equivalently, number of Trotter steps P). The protocols have been scaled according to Eq. (4.26) with $\alpha = 1.75$. The inset shows the unscaled protocols.

standard route is that of a linear-schedule continuous-time QA, henceforth referred to as "linear-QA", where $s(t) = t/\tau$. As we saw in Chapter (3) is well known [Dziarmaga, 2005; Zurek *et al.*, 2005] to lead to a Kibble-Zurek (KZ) [Kibble, 1976; Zurek, 1985; Polkovnikov *et al.*, 2011] power-law residual energy $\epsilon^{\text{res}}(\tau) \sim \tau^{-1/2}$. A closely related method involves the digitalization of QA, for instance with $\Delta t_m = 1$ (in units of $\hbar/J$). Figure 4.8 shows that both linear-QA and linear-dQA display the correct KZ behaviour $\epsilon^{\text{res}}(\tau) \sim \tau^{-1/2}$, with dQA only separated by a constant offset due to the discretization error.

Next, we consider other optimized schedules that have been proposed in the context of continuous-time QA. One, which we already discussed in Chapter 1, is the schedule proposed by Roland and Cerf [2002] where $s(t)$ has the form:

$$s = \frac{1}{2} + \frac{\tan(C(2t/\tau - 1))}{2\tan C} \tag{4.32}$$

where C is a parameter determining the slope at the critical point $s_c = 1/2$. Since we are working in the infinite chain limit, the system is gapless so C depends on τ and should be optimized. Alternatively, one can consider a power-law schedule, as proposed in [Barankov and Polkovnikov, 2008]:

$$s(t) = \frac{1}{2} + \frac{1}{2}\text{sgn}\left(\frac{t}{\tau} - \frac{1}{2}\right)\left|2\frac{t}{\tau} - 1\right|^C \tag{4.33}$$

C being now the power-law exponent, again dependent in general on τ and to be optimized. Both these strategies exploit the knowledge of the critical point location, here at $s_c = 1/2$, and can be applied either within a continuous-time QA, or, after digitalization, as dQA. Numerically, they both produce an improvement over linear-QA, with $\epsilon^{\text{res}} \sim \tau^{-\alpha}$, where $\alpha \sim 0.75$ and $\alpha \sim 0.8$ (for $\tau \lesssim 250\frac{\hbar}{J}$). In all cases, the digitalization appears to add a constant offset *upwards* to the continuous-time curves, with identical power-law exponent, confirming what we saw for linear dQA in Chapter 3. This, again, seems to be at variance with what the Trotter error does in Simulated Path-Integral Monte Carlo QA [Santoro *et al.*, 2002; Heim *et al.*, 2015; Mbeng *et al.*, 2019].

Finally, Fig. 4.8 shows the residual energy corresponding to the optimal *digitized*-QA solution, with τ calculated from (4.4). Here the behaviour of $\epsilon^{\text{res}}(\tau)$ shows the optimal power-law $\epsilon^{\text{res}} \sim \tau^{-1}$, coherently with the bound $\epsilon^{\text{res}}_{\text{P}} \geq (2\text{P} + 2)^{-1}$ and with $\tau \propto \text{P}$.

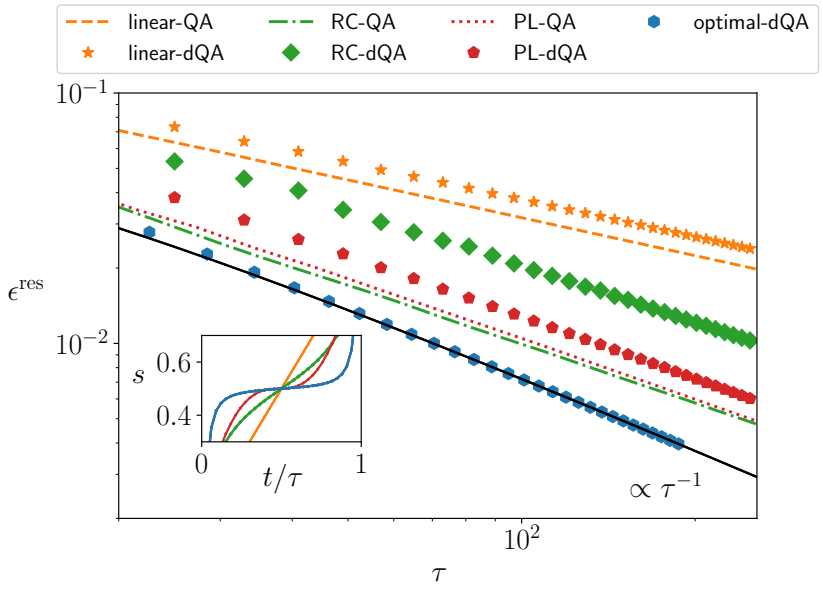

FIGURE 4.8: Scaling of the residual energy for various QA schedules applied to the infinite Ising chain problem. All *digitized*-QA data assume a Trotter discretization with $\Delta t_m = 1$. The linear continuous-time QA (orange solid solid line) and *digitized*-QA (orange stars) show a Kibble-Zurek exponent $\epsilon^{\rm res} \sim \tau^{-1/2}$. The Roland-Cerf QA and dQA (green dash-dot line and diamonds) and the power-law QA and dQA (red dotted line and pentagons) with optimized parameters show $\epsilon^{\rm res} \sim \tau^{-\alpha}$ with $\alpha \approx 0.75$ and $\alpha \approx 0.8$, respectively. The brown triangles represent the optimal QAOA regular results. The inset shows the values of $s(t)$ for fixed $\tau = 32$ for the different schedules.

The regular optimal dQA solution has the best possible performance, saturating the residual energy bound: $\epsilon^{\rm res} \sim \tau^{-1}$. However, such a quadratic speed-up over the plain KZ exponent comes with an extra computational cost to find the global QAOA variational minimum. Figure 4.9 shows that the number of iterations $n_{\rm iter}$ to find a minimum by starting from a random initial point increases as P^2, while $n_{\rm iter} \propto \sqrt{\mathrm{P}}$ for the iterative search of the regular optimal dQA solution.

Let us estimate how the residual energy decreases as a function of the computational cost $t_{\rm cc}$. One of the issues is the computational cost associated with a call of the "quantum oracle". Suppose we agree that such a cost scales with P, the number of unitaries involved in preparing $|\psi_{\rm P}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$, so that $t_{\rm cc} \propto \mathrm{P} n_{\rm iter}$. Then, the linear-dQA has $t_{\rm cc} \propto \mathrm{P}$, the random search of the optimal solution has $t_{\rm cc} \propto n_{\rm iter}\,\mathrm{P} \propto \mathrm{P}^3$, and the iterative search of the optimal dQA solution has $t_{\rm cc} \propto n_{\rm iter}\,\mathrm{P} \propto \mathrm{P}^{3/2}$. Using these estimates, we can express the residual energies in

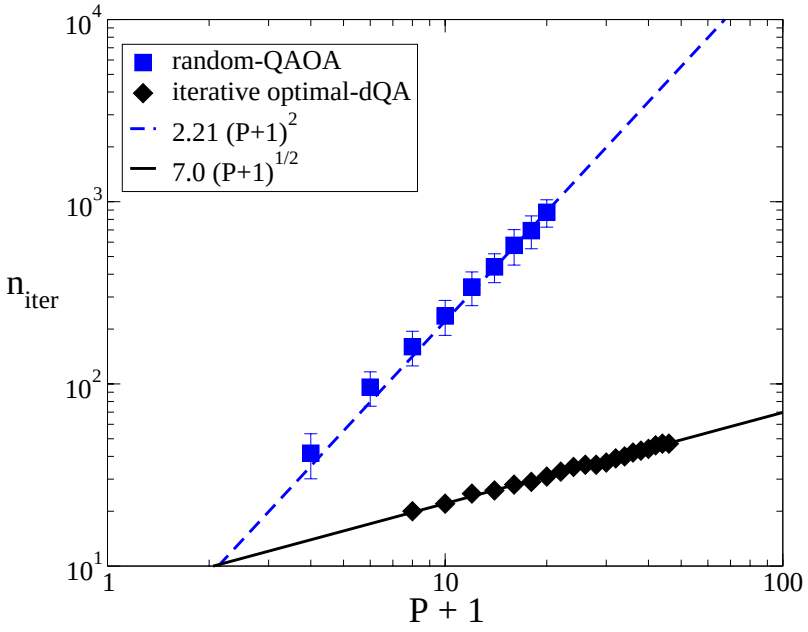

FIGURE 4.9: Number of iterations needed by the QOAO algorithm to converge to a minimum, with tolerance 10^{-5}. The black diamonds refer to the iterative search where the search system is initialized by interpolating a solution obtained for a lower value of P, while the blue squares represent a brute-force search starting from a random initial point. The classical optimization is performed using the BFGS algorithm [Nocedal and Wright, 2006].

terms of the computational cost:

$$\begin{cases} \epsilon^{\text{res}} \sim \mathrm{P}^{-\frac{1}{2}} \sim t_{\text{cc}}^{-\frac{1}{2}} & \text{(linear-dQA)} \\ \epsilon^{\text{res}} \sim \mathrm{P}^{-1} \sim t_{\text{cc}}^{-\frac{1}{3}} & \text{(QAOA, random)} \\ \epsilon^{\text{res}} \sim \mathrm{P}^{-1} \sim t_{\text{cc}}^{-\frac{2}{3}} & \text{(optimal-dQA, recursive)} \end{cases} . \tag{4.34}$$

Hence the overall performance of the optimal QAOA for a random initialization, in terms of computational time, is definitely *worse* than plain linear-dQA. To improve over linear-dQA, one *must* use a recursive initialization, leading to an optimal-dQA.

4.5 Summary and final remarks

We unveiled deep connections between Quantum Annealing (QA), in its digitized version, with the hybrid quantum-classical variational approach known as QAOA, elucidating the connection between optimal Quantum Control and the adiabaticity requirement of the driving protocols.

Two are the main results contained in this Chapter. The first is a technique to establish a variational bound on the residual energy of MaxCut problems on 2-regular periodic graphs by playing with the boundary conditions on the reduced spin problem. Such a technique can be naturally extended to higher-dimensional problems, and allows, through the use of Lieb-Robinson bounds, the physical evolution time to enter the game. We will deal with these issues in a Chapter 5, where we will discuss the presence of a light-cone, associated with a local Hamiltonian. In one dimension, we have shown that the variational bound $\epsilon_{\rm P}^{\rm res} \geq (2{\rm P}+2)^{-1}$ is precisely saturated by the Jordan-Wigner results, which also helps to elucidate the geometric nature of the minimization problem and the role of the variational parameters, 2P, in comparison with the number of spins, N. This, in turn, shows that the system becomes controllable, and the residual energy drops to 0, as soon as $2{\rm P} \geq N$.

The second important contribution has to do with the link between Quantum Control, which generally predicts the optimal schedule to be of the bang-bang form [Yang *et al.*, 2017], hence justifying the QAOA *Ansatz* [Farhi *et al.*, 2014], and the adiabatic dynamics behind QA, or more precisely here *digitized*-QA [Barends *et al.*, 2016]. Indeed, we have explicitly shown that among a large number of QAOA optimal solutions — $2^{\rm P}$ for $2{\rm P} < N$, a continuum for $2{\rm P} > N$ — one can iteratively single-out a *smooth regular* solution which can be regarded as the *optimal* digitized-QA schedule. Such a regular optimal solution provides a clear speed-up over linear-QA. The speed-up is quadratic — as in the Grover problem [Roland and Cerf, 2002] — if the computational cost for finding the solution is not considered. The speed-up still survives even when we account for the cost of searching the minimum, but only if smart iterative techniques [Pichler *et al.*, 2018] to construct the optimal solutions are used.

One point which is worth remarking is that the *smooth regular adiabatic digitized*-QA solution that we construct *does not use* any prior knowledge on the location of the critical point of the problem, nor any other spectral information, at variance with alternative schedule optimization approaches [Roland and Cerf, 2002; Barankov and Polkovnikov, 2008] which are explicitly tailored from the known critical bottleneck of the QA evolution.

In the next Chapter, we will generalize these results to the problem of preparing the ground state of more general quantum spin Hamiltonians in higher dimensions.

Chapter 5

Ground state preparation for the spin models

In this Chapter we generalize and further expand upon the ideas that were presented in Chapter 4. As before, our contribution can be conceptually divided in two parts, which are closely related to each other.

On the one hand, we extend the construction of *optimal digital*-QA schedules to the setting of quantum ground state preparation, where the target Hamiltonian $\widehat{H}_{\mathrm{T}}$ does not necessarily arise from a classical problem. We discuss how such a generalization is relevant, as it allows to use digital quantum devices to study low-energy proprieties of non-trivial quantum Hamiltonians which may exhibit exotic phases. Using numerical tools we explicitly construct the optimal schedule in the particular case of the XY model, where a size scaling is feasible.

On the other hand, we provide various generalizations of the variational bound for the residual energy which was derived, for the quantum Ising chain, in Chapter 4. As illustrative examples of this procedure, we consider higher dimensional spin systems (Ising model in arbitrary dimensions), classical disordered systems (random Ising chain) and quantum XY model. Finally, we go beyond the *digitized*-QA paradigm and give an extension of the variational bound in the framework of continuous-time evolution, through the use of Lieb-Robinson bounds.

The Chapter is organized as follows. In Sec. 5.1 we state the ground state preparation problem that we want to tackle. In Sec. 5.2 we discuss the problem of ground state preparation with digital circuits, we derive variational bounds on the residual energy for generic translational invariant systems and, in the benchmark case of the one-dimensional XY model, explicitly show how to use this framework

to construct *optimal digitized*-QA schedules. In Sec 5.4, by considering a random Ising chain, we show how to further extend the bounds to include disorder systems. In Sec. 5.5 we generalize the variational bounds to the case of continuous-time schedules. Finally, we compare our bounds for the QA of a classical Ising model with previously known results from Adiabatic Perturbation Theory (APT) [De Grandi and Polkovnikov, 2010; Polkovnikov, 2005].

5.1 Ground state preparation: statement of the problem

Let us consider spin-1/2 Hamiltonians defined on a regular lattice $\mathcal{L} = \{\mathbf{r}\}$ of N sites with periodic boundary conditions (PBC). Quite generally, the site vectors can be expressed as $\mathbf{r} = n_1\mathbf{a}_1 + \cdots + n_d\mathbf{a}_d$, where $\mathbf{a}_1 \cdots \mathbf{a}_d$ denote the d-dimensional primitive vectors of the lattice, and $n_i = 0, \cdots, L-1 \pmod L$ are integers, or equivalently, by the collection of the coordinates $\mathbf{n} = (n_1, \cdots, n_d)$. To simplify our notation, however, we will henceforth restrict our considerations to a two-dimensional square lattice with $N = L^2$ sites, where $\mathbf{a}_1 = \hat{\boldsymbol{x}}$ and $\mathbf{a}_2 = \hat{\boldsymbol{y}}$ (we assume the lattice constant $a = 1$) and $\mathbf{r} = n\hat{\boldsymbol{x}} + m\hat{\boldsymbol{y}}$, with $n, m = 0, \cdots, L-1 \pmod L$. We will equivalently denote as $\hat{\sigma}^{\alpha}_{\mathbf{r}} \longleftrightarrow \hat{\sigma}^{\alpha}_{n,m}$, with $\alpha = x, y, z$, Pauli matrices at site $\mathbf{r}$. The generalization to different lattices, and in higher dimension, is quite simple.

The spin Hamiltonian is assumed to depend linearly on $\mathrm{n_c}$ control parameters $\mathbf{g} = (g_1, \cdots, g_{\mathrm{n_c}})$:

$$\widehat{H}(\mathbf{g}) = g_1\widehat{H}_1 + \cdots + g_{\mathrm{n_c}}\widehat{H}_{\mathrm{n_c}} \ . \tag{5.1}$$

The hypothesis we will make for the different Hamiltonian terms are: *i)* locality, *ii)* translational invariance, and *iii)* symmetry under spin-flip, as implemented by the operator $\widehat{\mathcal{P}} = \prod_{\mathbf{r}} \hat{\sigma}^{x}_{\mathbf{r}}$.

Let us exemplify the terms we have in mind. A standard choice is the transverse field term:

$$\widehat{H}_x = -\sum_{\mathbf{r}} \hat{\sigma}^{x}_{\mathbf{r}} \ , \tag{5.2}$$

which we will always include as our first control Hamiltonian $\widehat{H}_1$. Other terms which could be considered are:

$$\begin{aligned} \widehat{H}_{\alpha\alpha,\mathbf{d}} &= \sum_{\mathbf{r}} \hat{\sigma}^{\alpha}_{\mathbf{r}} \hat{\sigma}^{\alpha}_{\mathbf{r}+\mathbf{d}} \\ \widehat{H}_{\alpha\alpha,\hat{\boldsymbol{y}}\pm\hat{\boldsymbol{x}}} &= \sum_{\mathbf{r}} \hat{\sigma}^{\alpha}_{\mathbf{r}} \hat{\sigma}^{\alpha}_{\mathbf{r}+\hat{\boldsymbol{y}}\pm\hat{\boldsymbol{x}}} \ , \end{aligned} \tag{5.3}$$

where $\alpha = x, y, z$ denotes directions in spin space, and $\mathbf{d} = \{\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}\}$ are nearest-neighbor vectors on the lattice. The first term would allow for *non-stoquastic* $\hat{\sigma}^x\hat{\sigma}^x$ or $\hat{\sigma}^y\hat{\sigma}^y$ driving and nearest-neighbor $\hat{\sigma}^z\hat{\sigma}^z$ coupling, while the second appears, for instance, in the frustrating next-nearest-neighbor term of the J_1J_2 Heisenberg model.

Notice that, as written, translational invariance is enforced. Rotational invariance is not required, neither on the lattice nor on the spin directions. Quite generally, we assume that for each term $\widehat{H}_{\mu=1\cdots n_c}$ included, there is a local operator $\widehat{O}_\mu(\mathbf{r})$ such that:

$$\widehat{H}_\mu = \sum_{\mathbf{r}\in\mathcal{L}} \widehat{O}_\mu(\mathbf{r}) = \sum_{\mathbf{r}\in\mathcal{L}} \widehat{T}^\dagger_{\mathbf{r}}\, \widehat{O}_\mu(\mathbf{r}_0)\, \widehat{T}_{\mathbf{r}} \ , \tag{5.4}$$

where $\widehat{T}_{\mathbf{r}}$ is the translation operator by a lattice vector $\mathbf{r}$

$$\widehat{T}^\dagger_{\mathbf{r}} \hat{\boldsymbol{\sigma}}_{\mathbf{r}_0} \widehat{T}_{\mathbf{r}} = \hat{\boldsymbol{\sigma}}_{\mathbf{r}_0+\mathbf{r}} \ , \tag{5.5}$$

and PBC are assumed. For the examples given above, we have $\widehat{O}_x(\mathbf{r}) = -\hat{\sigma}^x_{\mathbf{r}}$, $\widehat{O}_{xx,\hat{\boldsymbol{x}}}(\mathbf{r}) = \hat{\sigma}^x_{\mathbf{r}}\hat{\sigma}^x_{\mathbf{r}+\hat{\boldsymbol{x}}}$, etc.

Given a final target value of the parameter $\mathbf{g}_{\mathrm{T}}$, our goal is to prepare the corresponding ground state $|\psi_{\mathrm{gs}}(\mathbf{g}_{\mathrm{T}})\rangle$ of $\widehat{H}(\mathbf{g}_{\mathrm{T}})$, starting, for simplicity, from $|\psi_0\rangle = |+\rangle^{\otimes N}$ where $|+\rangle$ is the eigenstate of $\hat{\sigma}^x$ with eigenvalue $+1$. The goal can be achieved by adopting different protocols, either digital or in continuous time.

In a continuous-time framework, one might allow for an arbitrary variation $\mathbf{g}(t)$ as a curve in parameter space connecting $\mathbf{g}_{\mathrm{in}}$ to $\mathbf{g}_{\mathrm{T}}$ in a given total time τ, corresponding to a time-dependent Hamiltonian:

$$\widehat{H}(\mathbf{g}(t)) = \sum_{\mu=1}^{n_c} g_\mu(t) \widehat{H}_\mu \ . \tag{5.6}$$

Starting for instance from the ground state of $\widehat{H}_x$

$$|\psi_0\rangle = |+\rangle^{\otimes N} = 2^{-N/2} \left(|\uparrow\rangle + |\downarrow\rangle\right)^{\otimes N} \tag{5.7}$$

the state of the system at time t is given by the unitary Schrödinger evolution $|\psi(t)\rangle = \widehat{U}_{\rm QA}(t,0)|\psi_0\rangle$ where

$$\widehat{U}_{\rm QA}(t,0) = \mathrm{Texp}\left(-\frac{i}{\hbar}\int_0^t \mathrm{d}t'\, \widehat{H}(\mathbf{g}(t'))\right) . \tag{5.8}$$

This is, in essence, a rather general form of Quantum Annealing, alias Adiabatic Quantum Computing.

To quantify the "distance" from the target ground state, several possibilities are given. In principle one might maximize the fidelity

$$\mathcal{F}_\tau[\mathbf{g}(t)] = \left|\langle\psi_{\rm gs}(\mathbf{g}_{\rm T})|\psi(\tau)\rangle\right|^2 , \tag{5.9}$$

or minimize the expectation value of $\widehat{H}(\mathbf{g}_{\rm T})$:

$$E_\tau[\mathbf{g}(t)] = \langle\psi(\tau)|\widehat{H}(\mathbf{g}_{\rm T})|\psi(\tau)\rangle . \tag{5.10}$$

In both cases, the notation underlines the fact that these are functionals of the curve $\mathbf{g}(t)$. Incidentally, these possibilities can be unified into the minimization of the expectation of an Hermitean operator $\langle\psi(\tau)|\hat{D}|\psi(\tau)\rangle$, with $\hat{D} = -|\psi_{\rm gs}(\mathbf{g}_{\rm T})\rangle\langle\psi_{\rm gs}(\mathbf{g}_{\rm T})|$ in the first case, and $\hat{D} = \widehat{H}(\mathbf{g}_{\rm T})$ in the second case. Notice, however, that the first approach requires a knowledge of the target state.

We now start our discussion of bounds to the ground state preparation from the digital case, leaving the continuous-time generalization to Sec. 5.5

5.2 Digital protocol for ground state preparation

It is convenient to start our discussion from the digital case, where a convenient protocol, known as VQCS [Ho and Hsieh, 2019] — essentially related to QAOA *Ansatz* [Farhi *et al.*, 2014] — can be written as:

$$|\psi_{\rm P}(\boldsymbol{\gamma}_1\cdots\boldsymbol{\gamma}_{\rm P})\rangle = \widehat{U}(\boldsymbol{\gamma}_{\rm P})\cdots\widehat{U}(\boldsymbol{\gamma}_1)|\psi_0\rangle \tag{5.11}$$

where

$$\widehat{U}(\boldsymbol{\gamma}_m) = \mathrm{e}^{-i\gamma_{1,m}\widehat{H}_1}\cdots\mathrm{e}^{-i\gamma_{\rm n_c},m\widehat{H}_{\rm n_c}} \tag{5.12}$$

and $\boldsymbol{\gamma}_m = (\gamma_{1,m},\cdots,\gamma_{\rm n_c,m})$ is an $\rm n_c$-dimensional vector of parameters, with $m = 1\cdots\rm P$. To shorten our notation we will simply denote by $\boldsymbol{\gamma}$ the collection of vectors $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{\rm P})$. On the variational state $|\psi_{\rm P}(\boldsymbol{\gamma})\rangle$ we can calculate the average

of the target Hamiltonian, obtaining the variational energy:

$$E_{\rm P}(\boldsymbol{\gamma};\mathbf{g}_{\rm T}) = \langle\psi_{\rm P}(\boldsymbol{\gamma})|\widehat{H}(\mathbf{g}_{\rm T})|\psi_{\rm P}(\boldsymbol{\gamma})\rangle \ . \tag{5.13}$$

To quantify the difference from the target ground state energy $E_{\rm gs}(\mathbf{g}_{\rm T})$ we use the residual energy, which, as we did in Eq. 4.7, we write in the conveniently rescaled form:

$$\epsilon^{\rm res}_{\rm P}(\boldsymbol{\gamma};\mathbf{g}_{\rm T}) \ = \ \frac{E_{\rm P}(\boldsymbol{\gamma};\mathbf{g}_{\rm T}) - E_{\rm gs}(\mathbf{g}_{\rm T})}{E_{\rm max}(\mathbf{g}_{\rm T}) - E_{\rm gs}(\mathbf{g}_{\rm T})} = \frac{e_{\rm P}(\boldsymbol{\gamma};\mathbf{g}_{\rm T}) - e_{\rm gs}(\mathbf{g}_{\rm T})}{e_{\rm max}(\mathbf{g}_{\rm T}) - e_{\rm gs}(\mathbf{g}_{\rm T})} \ , \tag{5.14}$$

where $e_{\rm gs} = E_{\rm gs}/N$, $e_{\rm P} = E_{\rm P}/N$ and $e_{\rm max} = E_{\rm max}/N$ are energy densities.

To illustrate the flexibility of the scheme, consider, for instance, targeting the ground state of the $\mathrm{J_1J_2}$ Heisenberg model

$$\begin{aligned}\widehat{H}_{\mathrm{J_1J_2}} \ &= \ \mathrm{J_1}\sum_{\mathbf{r}}\left(\hat{\boldsymbol{\sigma}}_{\mathbf{r}}\cdot\hat{\boldsymbol{\sigma}}_{\mathbf{r}+\hat{\boldsymbol{x}}} + \hat{\boldsymbol{\sigma}}_{\mathbf{r}}\cdot\hat{\boldsymbol{\sigma}}_{\mathbf{r}+\hat{\boldsymbol{y}}}\right) \\ &+ \ \mathrm{J_2}\sum_{\mathbf{r}}\left(\hat{\boldsymbol{\sigma}}_{\mathbf{r}}\cdot\hat{\boldsymbol{\sigma}}_{\mathbf{r}+\hat{\boldsymbol{y}}+\hat{\boldsymbol{x}}} + \hat{\boldsymbol{\sigma}}_{\mathbf{r}}\cdot\hat{\boldsymbol{\sigma}}_{\mathbf{r}+\hat{\boldsymbol{y}}-\hat{\boldsymbol{x}}}\right) \ ,\end{aligned} \tag{5.15}$$

which possesses, depending on $\mathrm{J_2/J_1}$, both Néel [Calandra Buonaura and Sorella, 1998; Sandvik, 1997] as well as potentially non-trivial spin liquid ground states [Capriotti *et al.*, 2001; Choo *et al.*, 2019]. We might set up the problem with $\mathrm{n_c} = 7$, by choosing:

$$\begin{aligned}\widehat{H}_1 \ &= \ \widehat{H}_x \ , \\ \widehat{H}_{2\cdots4} \ &= \ \widehat{H}_{xx\cdots zz,\hat{\boldsymbol{x}}} + \widehat{H}_{xx\cdots zz,\hat{\boldsymbol{y}}} \ , \\ \widehat{H}_{5\cdots7} \ &= \ \widehat{H}_{xx\cdots zz,\hat{\boldsymbol{y}}+\hat{\boldsymbol{x}}} + \widehat{H}_{xx\cdots zz,\hat{\boldsymbol{y}}-\hat{\boldsymbol{x}}} \ .\end{aligned} \tag{5.16}$$

The target ground state we seek is therefore given by $\mathbf{g}_{\rm T} = (0, \mathrm{J_1}, \mathrm{J_1}, \mathrm{J_1}, \mathrm{J_2}, \mathrm{J_2}, \mathrm{J_2})$. As for the driving terms, we might include any of the previous Hamiltonian terms, including a proper subset, setting, for instance, $\gamma_{5\cdots7} \equiv 0$: the latter choice amounts to driving the system with a transverse field plus nearest-neighbor coupling terms only.

5.2.1 Residual energy bound for digital preparation

We will now derive a residual energy bound by following the technique introduced in Chapter 4. As prescribed by QAOA and VQCS we assume that each of the allowed driving Hamiltonians $\widehat{H}_\mu$ is a sum of commuting operators so that

$[\widehat{O}_\mu(\mathbf{r}_1), \widehat{O}_\mu(\mathbf{r}_2)] = 0$. The translational invariance of the various Hamiltonian terms $\widehat{H}_\mu$ and of the initial state $|\psi_0\rangle$ implies that we can write $E_{\rm P}(\boldsymbol{\gamma}; \mathbf{g}_{\rm T})$ as:

$$E_{\rm P}(\boldsymbol{\gamma}; \mathbf{g}_{\rm T}) = N \sum_{\mu=1}^{\rm n_c} g_\mu \langle \psi_{\rm P}(\boldsymbol{\gamma}) | \widehat{O}_\mu(\mathbf{r}_0) | \psi_{\rm P}(\boldsymbol{\gamma}) \rangle \ , \tag{5.17}$$

where $\mathbf{r}_0$ is any site of the lattice. As argued in Chapter 4, the application of the digital unitary operator $\widehat{U}_{\rm digit}(\boldsymbol{\gamma}) = \widehat{U}(\boldsymbol{\gamma}_{\rm P}) \cdots \widehat{U}(\boldsymbol{\gamma}_1)$ to any local operator $\widehat{O}_\mu(\mathbf{r}_0)$ involves only spins which have a distance which is a simple linear function of P from the site $\mathbf{r}_0$. For instance, if we use only single-spin flip and nearest-neighbor $\sum_{\mathbf{d}} \hat{\sigma}^z_{\mathbf{r}} \hat{\sigma}^z_{\mathbf{r}+\mathbf{d}}$ terms (with $\mathbf{d} = \hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}$), then $L_{\rm R} = 2{\rm P} + 2$ sites are involved in each direction, leading to a reduced spin problem on a lattice $\mathcal{L}_{\rm R}$ with $N_{\rm R} = L_{\rm R}^2$ sites. If we also include a double-spin-flip term $\sum_{\mathbf{d}} \hat{\sigma}^x_{\mathbf{r}} \hat{\sigma}^x_{\mathbf{x}+\mathbf{d}}$, then $L_{\rm R} = 4{\rm P} + 2$, etc. Notice that this applies as long as $L_{\rm R} \leq L$, while for larger P all spins are involved, and one should set $L_{\rm R} \equiv L$ and enforce PBC.

The crucial property which we will now use is, as discussed in Chapter 4, the freedom in the boundary conditions applied to the reduced spin problem. Indeed, for $L_{\rm R} \leq L$ the interaction terms involved in the calculation of the terms

$$\langle \psi_0 | \widehat{U}^\dagger_{\rm digit}(\boldsymbol{\gamma}) \widehat{O}_\mu(\mathbf{r}_0) \widehat{U}_{\rm digit}(\boldsymbol{\gamma}) | \psi_0 \rangle \tag{5.18}$$

never involve boundary couplings between spins, and there is therefore a certain freedom in the boundary conditions used. We will exploit this freedom by imposing either PBC or anti-periodic boundary conditions (ABC) in one or both directions.

To do so, it is convenient to define modified translation operators along the two directions as follows:

$$\widehat{T}^{(-)}_{\hat{\boldsymbol{x}}} = \widehat{T}_{\hat{\boldsymbol{x}}} \prod_{m=0}^{L_{\rm R}-1} \hat{\sigma}^x_{0,m} \ , \quad \widehat{T}^{(-)}_{\hat{\boldsymbol{y}}} = \widehat{T}_{\hat{\boldsymbol{y}}} \prod_{n=0}^{L_{\rm R}-1} \hat{\sigma}^x_{n,0} \ . \tag{5.19}$$

In words, $\widehat{T}^{(-)}_{\hat{\boldsymbol{x}}}$ flips all spins in the first column, and translates by one lattice spacing along $\hat{\boldsymbol{x}}$, and similarly for $\widehat{T}^{(-)}_{\hat{\boldsymbol{y}}}$, with a flip of the bottom row. Straightforward spin algebra shows that:

$$\widehat{T}^{(-)\dagger}_{\hat{\boldsymbol{x}}} \hat{\sigma}^\alpha_{n,m} \widehat{T}^{(-)}_{\hat{\boldsymbol{x}}} = \begin{cases} \hat{\sigma}^\alpha_{n+1,m} & \text{if } n < L_{\rm R} - 1 \\ \hat{\sigma}^x_{0,m} \hat{\sigma}^\alpha_{0,m} \hat{\sigma}^x_{0,m} & \text{if } n = L_{\rm R} - 1 \end{cases} \tag{5.20}$$

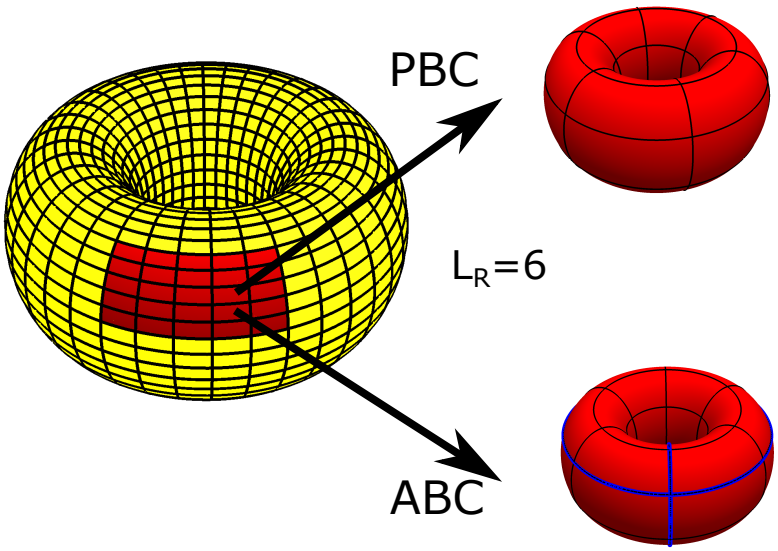

FIGURE 5.1: Illustration of the reduced spin lattice. An $L_{\mathrm{R}} \times L_{\mathrm{R}}$ piece of the original lattice with PBC (here $L_{\mathrm{R}} = 6$, in red) is singled-out and periodically closed either with PBC or with ABC. The thick blue lines denote the frustrating boundaries in the ABC case.

and similarly

$$\widehat{T}^{(-)\dagger}_{\hat{\boldsymbol{y}}} \hat{\sigma}^{\alpha}_{n,m} \widehat{T}^{(-)}_{\hat{\boldsymbol{y}}} = \begin{cases} \hat{\sigma}^{\alpha}_{n,m+1} & \text{if } m < L_{\mathrm{R}} - 1 \\ \hat{\sigma}^{x}_{n,0} \hat{\sigma}^{\alpha}_{n,0} \hat{\sigma}^{x}_{n,0} & \text{if } m = L_{\mathrm{R}} - 1 \end{cases} \tag{5.21}$$

Observe that since

$$\hat{\sigma}^{x} \begin{pmatrix} \hat{\sigma}^{x} \\ \hat{\sigma}^{y} \\ \hat{\sigma}^{z} \end{pmatrix} \hat{\sigma}^{x} = \begin{pmatrix} \hat{\sigma}^{x} \\ -\hat{\sigma}^{y} \\ -\hat{\sigma}^{z} \end{pmatrix} \tag{5.22}$$

this effectively introduces a minus sign in all the $\hat{\sigma}^{y}$ and $\hat{\sigma}^{z}$ terms appearing on the boundary sites. For all other sites, $\widehat{T}^{(-)}$ acts as an ordinary translation operator.

For ease of notation, we will henceforth denote by $\widehat{T}^{(+)}$ the ordinary translation operator $\widehat{T}$. Interestingly, $\widehat{T}^{(-)}_{\hat{\boldsymbol{x}}}$ and $\widehat{T}^{(-)}_{\hat{\boldsymbol{y}}}$ commute among themselves and with the corresponding ordinary translations: $\widehat{T}^{(-)}_{\hat{\boldsymbol{x}}} \widehat{T}^{(-)}_{\hat{\boldsymbol{y}}} = \widehat{T}^{(-)}_{\hat{\boldsymbol{y}}} \widehat{T}^{(-)}_{\hat{\boldsymbol{x}}}$, $\widehat{T}^{(-)}_{\hat{\boldsymbol{x}}} \widehat{T}^{(+)}_{\hat{\boldsymbol{y}}} = \widehat{T}^{(+)}_{\hat{\boldsymbol{y}}} \widehat{T}^{(-)}_{\hat{\boldsymbol{x}}}$, $\widehat{T}^{(+)}_{\hat{\boldsymbol{x}}} \widehat{T}^{(-)}_{\hat{\boldsymbol{y}}} = \widehat{T}^{(-)}_{\hat{\boldsymbol{y}}} \widehat{T}^{(+)}_{\hat{\boldsymbol{x}}}$. A straightforward calculation also shows that:

$$\begin{aligned} \widehat{T}^{(-)}_{n\hat{\boldsymbol{x}}} \overset{\text{def}}{=} \left(\widehat{T}^{(-)}_{\hat{\boldsymbol{x}}} \right)^{n} &= \widehat{T}^{(+)}_{n\hat{\boldsymbol{x}}} \prod_{m=0}^{L_{\mathrm{R}}-1} \prod_{i=0}^{n-1} \hat{\sigma}^{x}_{i,m} \\ \widehat{T}^{(-)}_{m\hat{\boldsymbol{y}}} \overset{\text{def}}{=} \left(\widehat{T}^{(-)}_{\hat{\boldsymbol{y}}} \right)^{m} &= \widehat{T}^{(+)}_{m\hat{\boldsymbol{y}}} \prod_{n=0}^{L_{\mathrm{R}}-1} \prod_{j=0}^{m-1} \hat{\sigma}^{x}_{n,j} \ , \end{aligned} \tag{5.23}$$

which in turn implies that

$$\widehat{T}^{(-)}_{L_{\rm R}\hat{\boldsymbol{x}}} = \widehat{T}^{(-)}_{L_{\rm R}\hat{\boldsymbol{y}}} = \widehat{\mathcal{P}} \,, \tag{5.24}$$

since $\widehat{T}^{(+)}_{L_{\rm R}\hat{\boldsymbol{x}}} = \widehat{T}^{(+)}_{L_{\rm R}\hat{\boldsymbol{y}}} = 1$.

Denoting by $s_x = \pm$ the choice of PBC/ABC in the $\hat{\boldsymbol{x}}$-direction, and similarly for $s_y = \pm$ in the $\hat{\boldsymbol{y}}$-direction, we can now define generalized translations of the form:

$$\widehat{T}^{(s_x s_y)}_{\mathbf{r}} = \widehat{T}^{(s_x s_y)}_{n\hat{\boldsymbol{x}}+m\hat{\boldsymbol{y}}} = \left(\widehat{T}^{(s_x)}_{\hat{\boldsymbol{x}}}\right)^n \left(\widehat{T}^{(s_y)}_{\hat{\boldsymbol{y}}}\right)^m \,, \tag{5.25}$$

where, evidently, $\widehat{T}^{(++)}_{n\hat{\boldsymbol{x}}+m\hat{\boldsymbol{y}}} = \widehat{T}_{n\hat{\boldsymbol{x}}+m\hat{\boldsymbol{y}}}$. Correspondingly, we can define Hamiltonians with any of the 4 choices of BC:

$$\hat{\mathcal{H}}^{(s_x s_y)}_{\mu} = \sum_{\mathbf{r}\in\mathcal{L}_{\rm R}} \widehat{T}^{(s_x s_y)\dagger}_{\mathbf{r}}\, \widehat{O}_{\mu}(\mathbf{0})\, \widehat{T}^{(s_x s_y)}_{\mathbf{r}} \,. \tag{5.26}$$

which have the important property of commuting with $\widehat{T}^{(s_x s_y)}_{\mathbf{r}}$:

$$\widehat{T}^{(s_x s_y)\dagger}_{\mathbf{r}}\, \hat{\mathcal{H}}^{(s_x s_y)}_{\mu}\, \widehat{T}^{(s_x s_y)}_{\mathbf{r}} = \hat{\mathcal{H}}^{(s_x s_y)}_{\mu} \,, \tag{5.27}$$

provided the corresponding $\widehat{O}_{\mu}$ is spin-flip symmetric, i.e.:

$$\widehat{\mathcal{P}}^{\dagger}\widehat{O}_{\mu}\widehat{\mathcal{P}} = \widehat{O}_{\mu} \,. \tag{5.28}$$

The elementary proof of these statements is given in Appendix C.

Let us now construct corresponding states in the Hilbert space of the reduced spin problem as follows:

$$|\widetilde{\psi}^{(s_x s_y)}_{\rm P}(\boldsymbol{\gamma})\rangle = \widehat{U}^{(s_x s_y)}(\boldsymbol{\gamma}_{\rm P}) \cdots \widehat{U}^{(s_x s_y)}(\boldsymbol{\gamma}_1)\, |\widetilde{\psi}_0\rangle \,, \tag{5.29}$$

with $|\widetilde{\psi}_0\rangle = |+\rangle^{\otimes N_{\rm R}}$ and the reduced evolution operators given by:

$$\widehat{U}^{(s_x s_y)}(\boldsymbol{\gamma}_m) = \mathrm{e}^{-i\gamma_{1,m}\hat{\mathcal{H}}^{(s_x s_y)}_1} \cdots \mathrm{e}^{-i\gamma_{{\rm n_c},m}\hat{\mathcal{H}}^{(s_x s_y)}_{\rm n_c}} \,. \tag{5.30}$$

We are now ready to state the main point behind the reduced spin problem. In essence, by an appropriate relabelling of the lattice sites we can always write:

$$\langle\psi_{\rm P}|\widehat{O}_{\mu}(\mathbf{r}_0)|\psi_{\rm P}\rangle = \langle\widetilde{\psi}^{(s_x s_y)}_{\rm P}|\widehat{O}_{\mu}(\mathbf{r}'_0)|\widetilde{\psi}^{(s_x s_y)}_{\rm P}\rangle \,, \tag{5.31}$$

where $\mathbf{r}'_0$ denotes the central bond of the reduced lattice, and we should notice that the right-hand side has the previously mentioned freedom on the boundary conditions.

To prove the promised bound, we will now use translational invariance for the constructed states and Hamiltonians. The fact that

$$\widehat{T}_{\mathbf{r}}^{(s_x s_y)} |\widetilde{\psi}_0\rangle = |\widetilde{\psi}_0\rangle \,, \tag{5.32}$$

since $|\widetilde{\psi}_0\rangle$ is translationally invariant and spin-flip symmetric by construction, and that $\hat{\mathcal{H}}_\mu^{(s_x s_y)}$ commutes with $\widehat{T}_{\mathbf{r}}^{(s_x s_y)}$, see Eq. (5.27), immediately imply the invariance of the reduced final state

$$\widehat{T}_{\mathbf{r}}^{(s_x s_y)} |\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}(\boldsymbol{\gamma})\rangle = |\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}(\boldsymbol{\gamma})\rangle \,. \tag{5.33}$$

Moreover, from Eqs. (5.27),(5.33) we deduce that:

$$\begin{aligned}
\langle\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}|\hat{\mathcal{H}}_\mu^{(s_x s_y)}|\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}\rangle &= \sum_{\mathbf{r}\in\mathcal{L}_{\mathrm{R}}} \langle\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}|\widehat{T}_{\mathbf{r}}^{(s_x s_y)\dagger}\, \widehat{O}_\mu(\mathbf{0})\, \widehat{T}_{\mathbf{r}}^{(s_x s_y)}|\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}\rangle \\
&= \sum_{\mathbf{r}\in\mathcal{L}_{\mathrm{R}}} \langle\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}|\widehat{O}_\mu(\mathbf{0})|\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}\rangle \\
&= N_{\mathrm{R}} \langle\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}|\widehat{O}_\mu(\mathbf{0})|\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}\rangle \\
&= N_{\mathrm{R}} \langle\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}|\widehat{O}_\mu(\mathbf{r}'_0)|\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}\rangle \,,
\end{aligned} \tag{5.34}$$

where the last step follows from the translational invariance of $|\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}\rangle$ under $\widehat{T}_{\mathbf{r}'_0}^{(s_x s_y)}$. Using Eqs. (5.17), (5.31), (5.34), and collecting together the different terms we conclude that:

$$\begin{aligned}
e_{\mathrm{P}}(\boldsymbol{\gamma};\mathbf{g}) &= \frac{1}{N_{\mathrm{R}}} \sum_{\mu=1}^{\mathrm{n_c}} g_\mu \langle\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}(\boldsymbol{\gamma})|\hat{\mathcal{H}}_\mu^{(s_x s_y)}|\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}(\boldsymbol{\gamma})\rangle \\
&= \frac{1}{N_{\mathrm{R}}} \langle\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}(\boldsymbol{\gamma})|\hat{\mathcal{H}}^{(s_x s_y)}(\mathbf{g})|\widetilde{\psi}_{\mathrm{P}}^{(s_x s_y)}(\boldsymbol{\gamma})\rangle \\
&\geq \max_{s_x,s_y} \frac{\mathcal{E}_{\mathrm{gs}}^{(s_x s_y)}(\mathbf{g})}{N_{\mathrm{R}}} \overset{\mathrm{def}}{=} \varepsilon_{\mathrm{R}}(\mathbf{g}) \,,
\end{aligned} \tag{5.35}$$

where the penultimate step follows from the standard variational principle for the reduced Hamiltonians, and the last defines the important quantity $\varepsilon_{\mathrm{R}}(\mathbf{g})$, which is variational lower bound for the final energy density. Notice that a dependence on P and on the number of controls $\mathrm{n_c}$ used is implicit in N_{R} and, as a consequence, in the ground state energy $\mathcal{E}_{\mathrm{gs}}^{(s_x s_y)}(\mathbf{g})$.

Inserting Eq. (5.35) in the expression for the residual energy we conclude that:

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma};\mathbf{g}_{\mathrm{T}}) \geq \epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}}) = \frac{\varepsilon_{\mathrm{R}}(\mathbf{g}_{\mathrm{T}}) - e_{\mathrm{gs}}(\mathbf{g}_{\mathrm{T}})}{e_{\mathrm{max}}(\mathbf{g}_{\mathrm{T}}) - e_{\mathrm{gs}}(\mathbf{g}_{\mathrm{T}})}, \tag{5.36}$$

where $\epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}})$ is the promised variational lower bound for the residual energy. We observe that $\varepsilon_{\mathrm{R}}(\mathbf{g}_{\mathrm{T}})$ essentially probes the sensitivity of the target state to the boundary conditions for a system of size N_{R}. Remarkably, this sets a lower bound for the residual energy *independently* of the intermediate digital dynamics that the system follows, as encoded in the parameters $\boldsymbol{\gamma}$.

5.3 Illustrative examples of digital preparation

We now apply the previous theory to tractable systems for which analytical or numerical results can be obtained. For illustration, we will consider three cases: 1) How to obtain the classical Néel antiferromagnetic state on a hyper-cubic lattice in d dimensions; 2) How to target the ground state of a transverse field Ising chain for a general value of the transverse field; 3) How to target the ground state of an XY model chain, essentially a generalized transverse field Ising chain.

5.3.1 The Néel state

Consider the Hamiltonian

$$\widehat{H}(\mathbf{g}) = g_1\widehat{H}_1 + g_2\widehat{H}_2 \tag{5.37}$$

where $\widehat{H}_1 = \widehat{H}_x$, and

$$\widehat{H}_2 = \sum_{\mathbf{r}}\sum_{i=1}^{d} \hat{\sigma}^z_{\mathbf{r}}\hat{\sigma}^z_{\mathbf{r}+\mathbf{a}_i} \tag{5.38}$$

is a classical antiferromagnetic Ising model on a d-dimensional cubic lattice of even system length L. The initial state $|\psi_0\rangle$ is the ground state corresponding to $\mathbf{g}_{\mathrm{in}} = (1,0)$, while we want to target the symmetric combination of the two classical Néel ground states of the Hamiltonian corresponding to $\mathbf{g}_{\mathrm{T}} = (0,1)$.

To compute the variational bound $\epsilon_{\mathrm{P}}^{\mathrm{var}}$, defined in Eq. (5.36), we need the quantities $e_{\mathrm{gs}}(\mathbf{g}_{\mathrm{T}})$, $e_{\mathrm{max}}(\mathbf{g}_{\mathrm{T}})$ and $\varepsilon_{\mathrm{R}}(\mathbf{g}_{\mathrm{T}})$. The Néel ground states and the fully magnetized state are respectively associated to the minimum and maximum energy densities $e_{\mathrm{gs}}(\mathbf{g}_{\mathrm{T}}) = -1$ and $e_{\mathrm{max}}(\mathbf{g}_{\mathrm{T}}) = 1$. We choose ABC on all the boundaries of the reduced problem so that the reduced ground state energy density is $\varepsilon_{\mathrm{R}}(\mathbf{g}_{\mathrm{T}}) =$

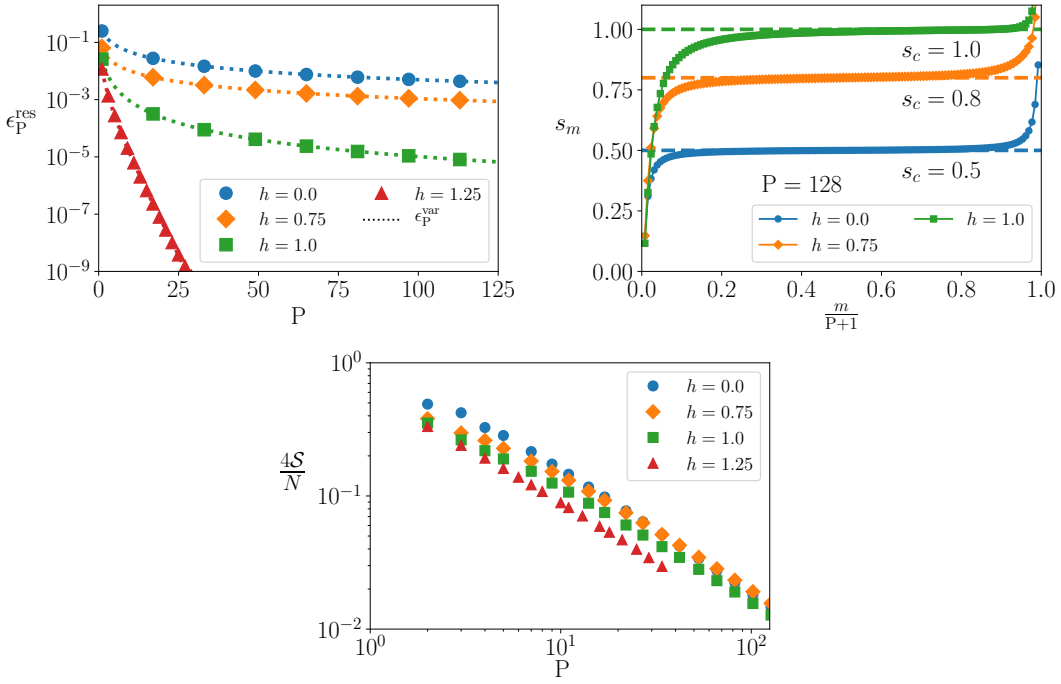

FIGURE 5.2: Digital optimal control protocol for various target h in the transverse field Ising chain of length $N = 1024$ (Eq. (5.43)). (a) Dependence on the circuit depth P of the optimal value of the residual energy $\epsilon_{\rm P}^{\rm res}$(full symbols) for various targets. Here the dotted lines represent the theoretical variational bound. (b) Regular optimal-dQA protocol for several targets. Here we plot $s_m = \gamma_{2,m}/(\gamma_{1,m} + (1-h)\gamma_{2,m})$ (full symbols). The dashed line represents the critical value $s_c = 1/(2-h)$. (c) Average Shannon entropy $\mathcal{S}$ defined in Eq. (5.49).

$-1 + \frac{2}{L_{\rm R}}$. We then get the bound:

$$\epsilon_{\rm P}^{\rm res}(\boldsymbol{\gamma}; \mathbf{g}_{\rm T}) \geq \frac{\varepsilon_{\rm R}(\mathbf{g}_{\rm T}) - e_{\rm gs}(\mathbf{g}_{\rm T})}{e_{\rm max}(\mathbf{g}_{\rm T}) - e_{\rm gs}(\mathbf{g}_{\rm T})} = \frac{1}{L_{\rm R}} = \frac{1}{2{\rm P}+2}\,, \tag{5.39}$$

where we explicitly used that the reduced problem's size is $L_{\rm R} = 2{\rm P} + 2 \leq L$. We note that the bound does not depend on the dimensionality of the system. It instead depends on $L_{\rm R}$ which can be increased by modifying the range of interactions.

5.3.2 Ising model in $d = 1$

We now consider the Transverse Field Ising Model (TFIM) in $d = 1$, with the driving Hamiltonians

$$\widehat{H}_1 = \widehat{H}_x = -\sum_j \hat{\sigma}_j^x \tag{5.40}$$

$$\widehat{H}_2 = \widehat{H}_{zz} = \sum_j \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z \, . \tag{5.41}$$

The discrete time evolution operator is then

$$\widehat{U}^{\mathrm{TFIM}}(\boldsymbol{\gamma}_m) = \mathrm{e}^{-i\gamma_{1,m}\widehat{H}_1}\mathrm{e}^{-i\gamma_{2,m}\widehat{H}_2} = \mathrm{e}^{-i\gamma_{1,m}\widehat{H}_x}\mathrm{e}^{-i\gamma_{2,m}\widehat{H}_{zz}} \, , \tag{5.42}$$

where $\boldsymbol{\gamma}_m$ is a vector of $\mathrm{n_c} = 2$ components. Starting from an initial state $\mathbf{g}_{\mathrm{in}} = (1, 0)$, and we target the ground state of the Hamiltonian corresponding to $\mathbf{g}_{\mathrm{T}} = (h, 1)$, i.e., to

$$\widehat{H}_{\mathrm{T}} = \sum_j \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z - h \sum_j \hat{\sigma}_j^x = \widehat{H}_{zz} + h\widehat{H}_x \, . \tag{5.43}$$

Using the Jordan Wigner transformation the problem can be mapped into a set of independent two level systems (or pseudo-spin) [Lieb *et al.*, 1961; Young, 1997]. In Appendix E, we use this framework to derive an analytical expression for the variational bound $\epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}})$ as a function of the target h and the circuit depth P. In Fig. 5.2(a) we show the dependence of $\epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}})$ (dotted lines) on the circuit depth. The asymptotic behavior of $\epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}})$ for deep circuits depends on the target h as follows:

$$\epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}}) \propto \begin{cases} \mathrm{P}^{-1} & \text{for } h < 1 \\ \mathrm{P}^{-2} & \text{for } h = 1 \\ \mathrm{e}^{-c\mathrm{P}} & \text{for } h > 1 \end{cases} \, . \tag{5.44}$$

Notice that the standard power-law P^{-1} for $h < 1$ changes into P^{-2} for $h = 1$ before turning into a decaying exponential for $h > 1$. Such a decaying exponential signals that the paramagnetic phase for $h > 1$ is gapped and *insensitive to the boundary conditions*. On the contrary, the gapped ferromagnetic phase for $h < 1$ is *not* insensitive to boundary conditions, due to the symmetry breaking.

Since $\epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}})$ is a rigorous variational bound, no schedule can achieve a residual energy lower than $\epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}})$. However, *a priori*, this statement is not informative on whether there exist a set of optimal variational parameters γ^{opt} that attain the

bound $\epsilon_{\rm P}^{\rm res}(\mathbf{g}_{\rm T}) = \epsilon_{\rm P}^{\rm var}(\mathbf{g}_{\rm T})$, nor it gives any indication on proprieties or patterns of $\gamma^{\rm opt}$.

Following Mbeng *et al.* [2019], we use a numerical iterative construction of $\gamma^{\rm opt}$ to address this issue.

We implement the function $E_{\rm P}(\boldsymbol{\gamma};\mathbf{g}_{\rm T})$ with `PyTorch` that provides built-in auto-differentiation routines [Paszke *et al.*, 2017]. We then search for a local minimum of $E_{\rm P}(\boldsymbol{\gamma};\mathbf{g}_{\rm T})$ with the Broyden-Fletcher-Goldfard-Shanno (BFGS) optimization algorithm [Nocedal and Wright, 2006], using back-propagation to compute the required gradients. Starting from P = 2, we proceed iteratively by using the results obtained at level P to build, trough interpolation, an initial guess for the solution at level 2P. When successful, this procedure speeds up the optimization process and can be used to find smooth solutions $\gamma^{\rm opt}$ to the optimization problem. In Fig. 5.2(a) we compare the residual energy $\epsilon_{\rm P}^{\rm res}(\boldsymbol{\gamma}^{\rm opt};\mathbf{g}_{\rm T})$ associated to the iteratively constructed smooth schedules (symbols) with the variational bound $\epsilon_{\rm P}^{\rm var}$ (dotted lines) for various targets. We find that the smooth solutions constructed iteratively satisfy the equality $\epsilon_{\rm P}^{\rm res}(\boldsymbol{\gamma}^{\rm opt};\mathbf{g}_{\rm T}) = \epsilon_{\rm P}^{\rm var}(\mathbf{g}_{\rm T})$ and are therefore optimal.

To address the issue of the adiabatic nature of this solution, let us introduce the intermediate state $|\psi_m(\boldsymbol{\gamma})\rangle$ after m digital steps, which satisfies the following discrete version of Schrödinger's equation

$$|\psi_0(\boldsymbol{\gamma})\rangle = |\psi_0\rangle \tag{5.45}$$

$$|\psi_{m+1}(\boldsymbol{\gamma})\rangle = \widehat{U}_m|\psi_m(\boldsymbol{\gamma})\rangle \tag{5.46}$$

where we recall that the discrete-time evolution operator is defined to be

$$\widehat{U}_m = \widehat{U}(\boldsymbol{\gamma}_m) = \mathrm{e}^{-i\gamma_{1,m}\widehat{H}_1}\cdots\mathrm{e}^{-i\gamma_{\mathrm{n_c},m}\widehat{H}_{\mathrm{n_c}}} \ . \tag{5.47}$$

We can always diagonalize $\widehat{U}_m$, finding an orthonormal basis $|\theta_m\rangle$:

$$\widehat{U}_m|\theta_m\rangle = \mathrm{e}^{-i\theta_m}|\theta_m\rangle \ . \tag{5.48}$$

We say that $\boldsymbol{\gamma}$ defines an adiabatic dynamics if the state $|\psi_m(\boldsymbol{\gamma})\rangle$ closely *follows an eigenstate* $|\overline{\theta}_m\rangle$ of $\widehat{U}_m$. This is a natural extension [Dranov *et al.*, 1998] of the concept of adiabaticity in continuous-time dynamics. Let $p_{\boldsymbol{\gamma}}(\theta_m) = |\langle\theta_m|\psi_m(\boldsymbol{\gamma})\rangle|^2$ be the probability of finding the system in a given eigenstate $|\theta_m\rangle$. The definition of adiabatic dynamics given above, suggests to quantify the degree of adiabaticy by measuring how close the distribution $p_{\boldsymbol{\gamma}}(\theta_m)$ is to a Kröneker-delta in θ_m. The

adiabaticity of the discrete dynamics with P steps, can then be quantified with the average Shannon entropy $\mathcal{S}_{\boldsymbol{\gamma}}(\mathrm{P})$ of the distribution $p_{\boldsymbol{\gamma}}(\theta_m)$:

$$\mathcal{S}_{\boldsymbol{\gamma}}(\mathrm{P}) = -\frac{1}{\mathrm{P}} \sum_{m=1}^{\mathrm{P}} \sum_{\theta_m} p_{\boldsymbol{\gamma}}(\theta_m) \log[p_{\boldsymbol{\gamma}}(\theta_m)] \; . \tag{5.49}$$

For an adiabatic dynamics $\mathcal{S}_{\boldsymbol{\gamma}}(\mathrm{P}) \to 0$ as $\mathrm{P} \to \infty$. Figure 5.2(c) shows that, indeed, the average Shannon entropy $\mathcal{S}$ computed on the instantaneous eigenbasis of the digital time-evolution operator vanishes as $\mathrm{P} \to \infty$. This signals that γ^{opt} induces a digital adiabatic dynamics in which the system remains close to an eigenstate of an instantaneous effective Hamiltonian. This concept is elaborated in more depth in App. D.

Since the solutions obtained iteratively have the property of being simultaneously smooth, optimal and adiabatic, in the following, we will refer to them as optimal digital Quantum Annealing (dQA) protocols γ^{dQA^*}. The usual interpolating strategy considered in AQC might help in visualizing the optimal dQA protocols. We write the interpolating Hamiltonian as

$$\widehat{H}(s) = (1-s)\widehat{H}_x + s\widehat{H}_{\mathrm{T}} = \Big(sh + (1-s)\Big)\widehat{H}_x + s\widehat{H}_{zz} \; . \tag{5.50}$$

Hence, after a lowest-order Trotter splitting, the parameters

$$\begin{cases} \gamma_{1,m} = \Big(hs_m + (1-s_m)\Big)\dfrac{\Delta t_m}{\hbar} \\ \gamma_{2,m} = s_m \dfrac{\Delta t_m}{\hbar} \end{cases} \tag{5.51}$$

would lead to an approximate evolution operator of the form:

$$\widehat{U}_{\mathrm{QA}}(\tau, 0) \approx \widehat{U}_{\mathrm{digit}}(\boldsymbol{\gamma}) = \widehat{U}(\boldsymbol{\gamma}_{\mathrm{P}}) \cdots \widehat{U}(\boldsymbol{\gamma}_1) \; . \tag{5.52}$$

Quite generally, given a set of parameters $\boldsymbol{\gamma}$ we can invert these equations to obtain the corresponding s_m and Δt_m. One finds:

$$s_m = \frac{\gamma_{2,m}}{\gamma_{1,m} + (1-h)\gamma_{2,m}} \; , \tag{5.53}$$

while

$$\frac{\Delta t_m}{\hbar} = \gamma_{1,m} + (1-h)\gamma_{2,m} \; . \tag{5.54}$$

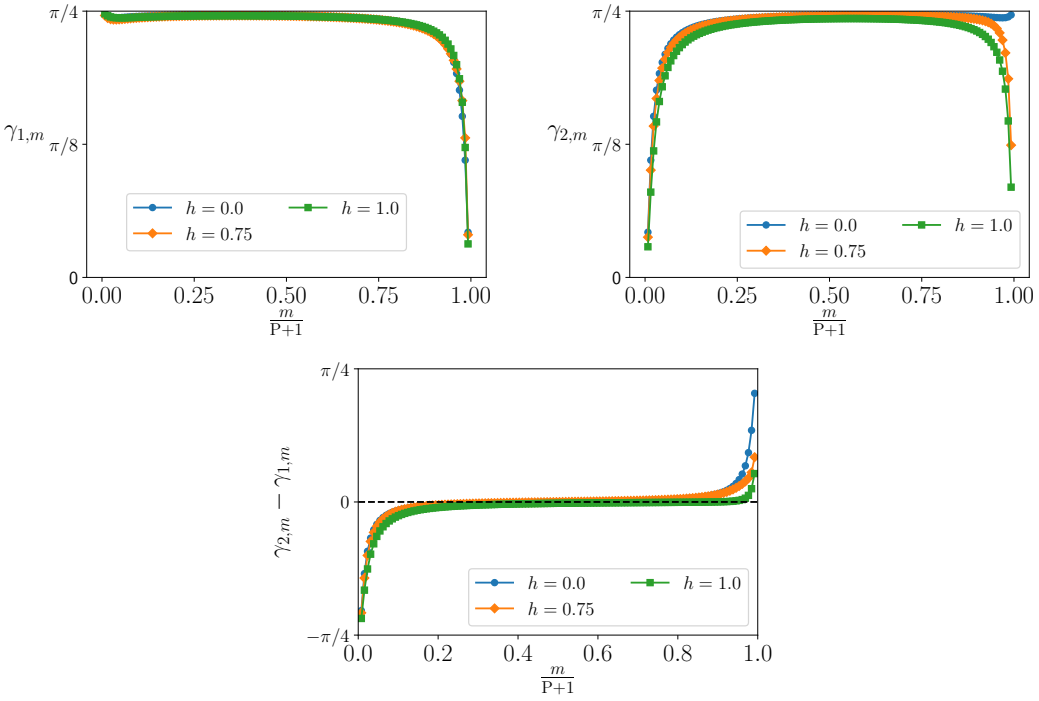

FIGURE 5.3: Optimal variational parameters for digital ground state preparation for the quantum Ising chain. The target fields considered are $h = 0, 0.5, 1$. We show $\gamma_{1,m}$ (a), $\gamma_{2,m}$ (b), and $\gamma_{2,m} - \gamma_{1,m}$ (c) as a function of m. Notice, in (c), the critical slowdown of the schedule in proximity of the "digital critical line" $\gamma_{2,m} - \gamma_{1,m} = 0$.

It is interesting to visualize the smooth optimal schedule s. It should show a flattening for

$$s_c = \frac{1}{2-h}\,, \tag{5.55}$$

where the gap closes, as long as $h \leq 1$. Fig. 5.2(b) shows how s_m, defined in Eq. (5.53), changes during the schedule for $h = 0$, $h = 0.75$, and $h = 1$. As expected, the optimal-dQA schedules drastically slowdown in proximity of the critical value s_c, which is a common feature of efficient adiabatic schedules [Roland and Cerf, 2002].

An alternative and more general visualization of the critical slowdown, can be given directly in terms of the optimal variational parameters $\boldsymbol{\gamma}^{\mathrm{dQA}^*}$. In Fig. 5.3 we show how $\gamma_{1,m}$ and $\gamma_{2,m}$ depend on m for the target fields $h = 0$, 0.5, 1. The variational parameter $\gamma_{1,m}$ depends only weakly on the target field h. In particular, for all values of h we find that $\gamma_{1,m}$ starts from $\gamma_{1,m=1} \approx \pi/4$ and smoothly varies to a final value $0 < \gamma_{1,m=\mathrm{P}} \lesssim 0.2$. For intermediated values of m, $\gamma_{1,m}$ stays close to $\pi/4$. Conversely, the variational parameter $\gamma_{2,m}$ exhibits a strong dependence on the target field h, particularly for $m \approx \mathrm{P}$. Indeed, the ground state of the

effective Hamiltonian, see App. D, at step $m = \mathrm{P}$, which depends on $\gamma_{1,\mathrm{P}}$ and $\gamma_{2,\mathrm{P}}$, should encode the target state.

Using $\gamma_{1,m}$ and $\gamma_{2,m}$ is also convenient to visualize the critical slowdown of the digital dynamics. In Appendix D, we show that the minimal gap Δ_m of the effective Hamiltonian that generates the digital evolution at step m is

$$\Delta_m = 2|\gamma_{2,m} - \gamma_{1,m}| + \mathcal{O}(|\gamma_{2,m} - \gamma_{1,m}|^2) + \mathcal{O}(N^{-1}) \, . \tag{5.56}$$

We therefore expect a slowdown of the adiabatic path in proximity of the vanishing gap, for $\gamma_{2,m} - \gamma_{1,m} = 0$, which is numerically observed, see Fig. 5.3(c).

Finally we point out that γ^{dQA^*} is not the only optimal solution. Indeed, there are 2^{P} distinct equivalent solution to the problem of minimizing $E(\boldsymbol{\gamma}; \mathbf{g}_{\mathrm{T}})$. However, these solution correspond to extremely different digital dynamics which are not adiabatic. The iterative construction we used to build γ^{dQA^*} is what enables us to single-out the adiabatic solution.

5.3.3 XY model in $d = 1$

As a generalization of the previous case, let us consider the problem of preparing the ground state of the one-dimensional XY model [Katsura, 1962; Mukherjee *et al.*, 2007]. A convenient two-dimensional parameterization of the target Hamiltonian in terms of two variables, the transverse field h and the anisotropy δ, is given by

$$\widehat{H}_{\mathrm{XY}} = -h \sum_j \hat{\sigma}_j^x + \frac{1-\delta}{2} \sum_j \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y + \frac{1+\delta}{2} \sum_j \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z \, . \tag{5.57}$$

Starting from the ground state of $\widehat{H}_x$ we now aim at preparing the ground state of $\widehat{H}_{\mathrm{XY}}$ at an arbitrary point in the (h, δ) phase diagram, shown in Fig. 5.4. The transverse field Ising model analyzed in the previous section corresponds to the line $\delta = 1$ (thick red line). The system is gapless along critical lines (thick black lines) $h = 1$, $h = -1$ and the segment $-1 \leq h \leq 1$ which lies on $\delta = 0$. In particular, when $h < 1$ an adiabatic schedule aiming at preparing the ground state of $\widehat{H}_{\mathrm{XY}}$ will necessarily encounter a vanishing gap.

We now discuss the choice of the driving Hamiltonian used to implement the dynamics. As a first guess, one would write the digital evolution operator as:

$$\widehat{U}^{\mathrm{XY}}(\boldsymbol{\gamma}_m) = \mathrm{e}^{-i\gamma_{1,m}\widehat{H}_x} \mathrm{e}^{-i\gamma_{2,m}\widehat{H}_{yy}} \mathrm{e}^{-i\gamma_{3,m}\widehat{H}_{zz}} \, , \tag{5.58}$$

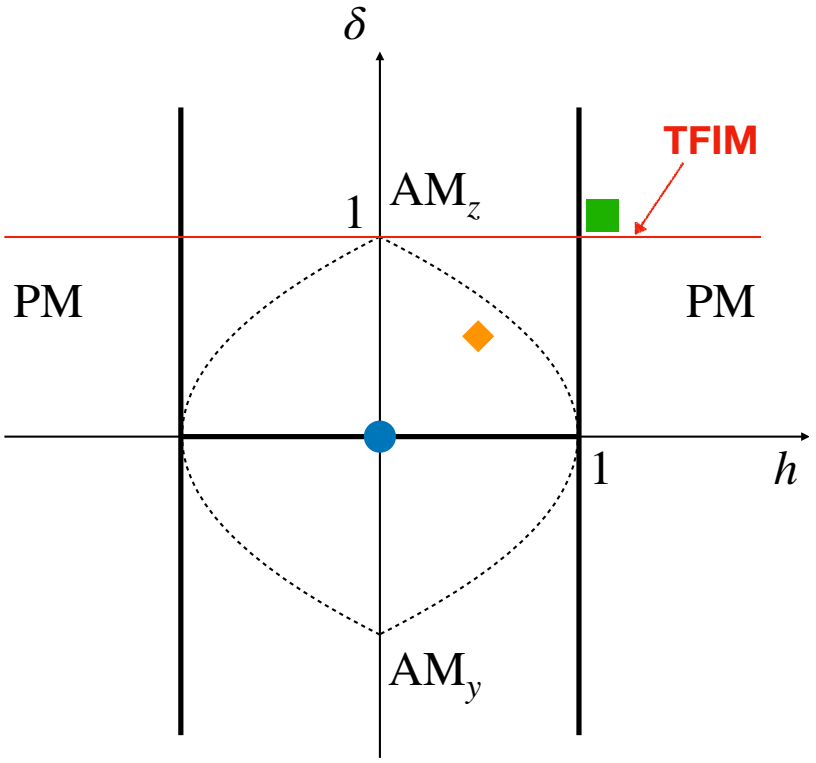

FIGURE 5.4: Phase diagram of the XY spin chain in Eq. (5.57): δ is the anisotropy parameter and h the transverse field. The three phases appearing are anti-ferromagnet along the y direction (AM$_y$), anti-ferromagnet along the z direction (AM$_z$) and paramagnet (PM). The different phases are separated by critical lines (thick black lines). The dashed line separates the commensurate and incommensurate anti-ferromagnetic phases [Mukherjee *et al.*, 2007].

where

$$\widehat{H}_{yy} = \sum_j \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y \,. \tag{5.59}$$

An immediate consideration which comes to mind is that such an approach appears redundant. Indeed, using the identity

$$\mathrm{e}^{-i\theta H_{yy}} = \mathrm{e}^{+i\frac{\pi}{4}\widehat{H}_x} \mathrm{e}^{-i\theta \widehat{H}_{zz}} \mathrm{e}^{-i\frac{\pi}{4}\widehat{H}_x} \,, \tag{5.60}$$

one would conclude that it should be possible to proceed with a digital evolution operator *identical* to that of the TFIM, i.e., involving only a repeated action of $\widehat{H}_x$ and $\widehat{H}_{zz}$:

$$\widehat{U}^{\mathrm{XY}}(\boldsymbol{\gamma}_m) = \mathrm{e}^{-i\tilde{\gamma}_{1,m}\widehat{H}_x} \mathrm{e}^{-i\tilde{\gamma}_{2,m}\widehat{H}_{zz}} \,. \tag{5.61}$$

This leads to an interesting observation: the effect of the $\widehat{H}_{yy}$ term, present in the Hamiltonian, can be effectively "engineered" in the dynamics by other terms, a fact that bears some similarities with the Floquet engineering techniques employed in

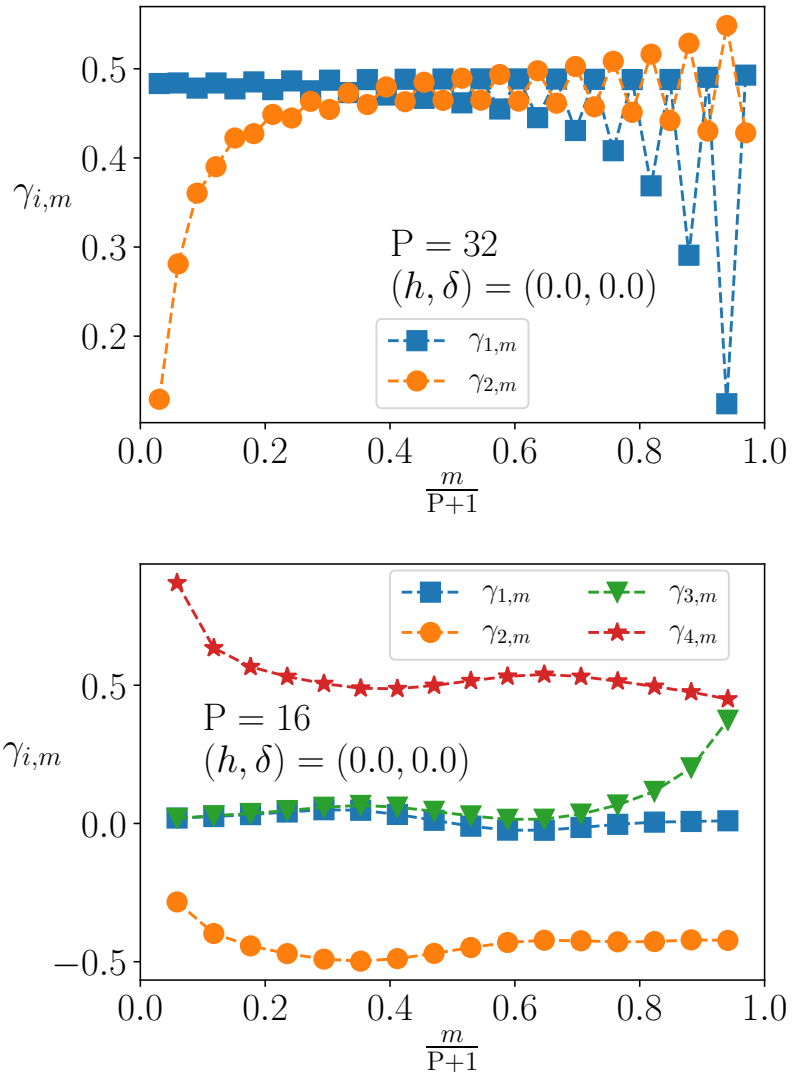

FIGURE 5.5: Optimal variational parameter for the digital ground state preparation of the point $(h = 0, \delta = 0)$ in the XY phase diagram, for a chain on $N = 1024$ sites. The optimal parameters where found with the iterative procedure described in the main text. (a) Optimal values of $\tilde{\gamma}_{1,m}$ and $\tilde{\gamma}_{2,m}$ for the digital dynamics is implemented by two driving Hamiltonians in Eq. (5.40), for P = 32. Notice that $\boldsymbol{\gamma}$ is not a smooth function of $\frac{m}{P+1}$. (b) Optimal values of $\gamma_{1,m}$, $\gamma_{2,m}$, $\gamma_{3,m}$ and $\gamma_{4,m}$ for the digital dynamics is implemented by four driving Hamiltonians in Eq. (5.62), for P = 16. In this case the $\boldsymbol{\gamma}$ a smooth function of $\frac{m}{P+1}$.

optimal control [Claeys *et al.*, 2019]. For our purpose, however, the procedure is rather unsatisfactory. Fig. 5.5(a) shows the optimal $\tilde{\gamma}_{i,m}$, as a function of m, when the digital dynamics is implemented as in Eq. (5.61): the optimal parameters are *not smooth* functions of m, as an evident even-odd staggering is visible.

It turns out that the ideal strategy is to write a digital evolution with four operators, as follows:

$$\widehat{U}^{\mathrm{XY}}(\boldsymbol{\gamma}_m) = \mathrm{e}^{-i\gamma_{1,m}\widehat{H}_x}\mathrm{e}^{-i\gamma_{2,m}\widehat{H}_{yy}}\mathrm{e}^{-i\gamma_{3,m}\widehat{H}_x}\mathrm{e}^{-i\gamma_{4,m}\widehat{H}_{zz}} \ , \tag{5.62}$$

which, by the identity in Eq. (5.60), can also be expressed as:

$$\widehat{U}^{\mathrm{XY}}(\boldsymbol{\gamma}_m) = \widehat{U}^{\mathrm{TFIM}}(\tilde{\boldsymbol{\gamma}}_{2m})\widehat{U}^{\mathrm{TFIM}}(\tilde{\boldsymbol{\gamma}}_{2m-1}) \ . \tag{5.63}$$

Here we have:

$$\boldsymbol{\gamma}_m = (\tilde{\gamma}_{1,2m} + \frac{\pi}{4}, \tilde{\gamma}_{2,2m}, \tilde{\gamma}_{1,2m-1} - \frac{\pi}{4}, \tilde{\gamma}_{2,2m-1}) \ . \tag{5.64}$$

Fig. 5.5(b) shows that when using the four driving Hamiltonians given in Eq. (5.62), the optimal parameters $\boldsymbol{\gamma}_m$ are smooth. We conclude that, although two driving Hamiltonians are sufficient to prepare any state, to realize an adiabatic dynamics, additional driving terms should be added. This shows that a digital dynamics of P steps with the XY driving Hamiltonians in Eq. (5.62) is equivalent to one of 2P steps with the TFIM driving Hamiltonians in Eq. (5.40). The number of parameters is clearly unchanged. With the latter four-term choice of digital dynamics, we have $\mathbf{g}_{\mathrm{in}} = (0, 0, 1, 0)$ and $\mathbf{g}_{\mathrm{T}} = (0, \frac{1-\delta}{2}, h, \frac{1+\delta}{2})$.

As in the TFIM, the Jordan-Wigner pseudo-spin decomposition allows us to find analytical expressions describing the variational bound $\epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}})$ on the target (h, δ). We illustrate the main features of the algorithm by considering three specific targets in the phase diagram in Fig 5.4. The three sample points considered correspond to $(h, \delta) = (0, 0)$ (blue circle), $(h, \delta) = (0.5, 0.5)$ (orange diamond) and $(h, \delta) = (1.1, 1.1)$ (green square). A more detailed analysis is presented in Appendix E. In Fig. 5.6 we show that, as in the TFIM, the variational bound is still saturated by the optimal-dQA schedule found with the iterative construction[1].

[1] In App. E we discuss the presence of negligible differences between $\epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}})$ and $\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma}^{\mathrm{dQA}^*}, \mathbf{g}_{\mathrm{T}})$ close to the line $\delta = 0$. However we also show that this discrepancy vanishes when P becomes sufficiently large.

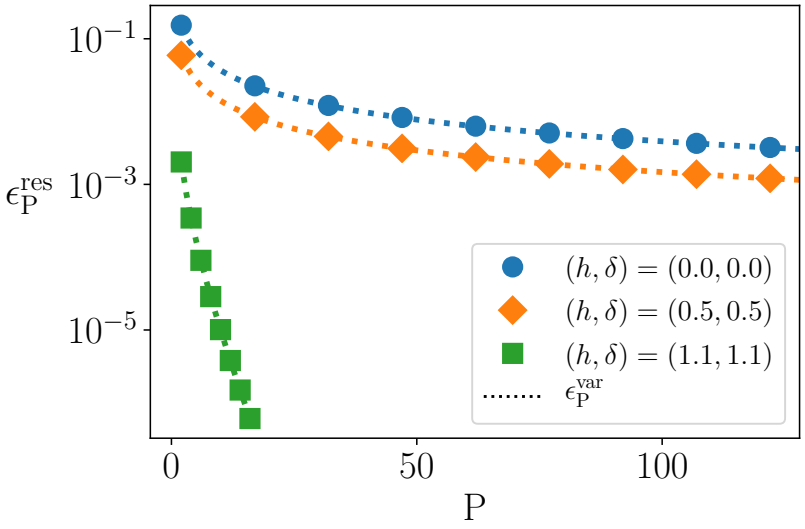

FIGURE 5.6: Digital ground state preparation for the XY model. Residual energy $\epsilon_{\mathrm{P}}^{\mathrm{res}}$ as a function of the depth P of the digital circuit. The symbols correspond to the residual energy associated to the optimal-dQA schedule, while the lines represent the theoretical variational bound derived in App. E.

Next, we consider the question of which states are easier to reach in the phase diagram. In the context of digital ground state preparation, it is natural to associate how easily a state can be reached at depth P, with the lowest residual energy achievable with a digital evolution of P steps. Intuitively, states that are easier to reach can be prepared with lower errors, as quantified by the residual energy. The situation is illustrated in Fig. 5.7, which shows that the hardest state to reach is the $(h = 0, \delta = 0)$ point, corresponding to a half-filled band of free fermions.

We now turn to a visualization of the optimal protocol $\boldsymbol{\gamma}^{\mathrm{dQA}^*}$. In App, D we show that the gap of the effective digital Hamiltonian is such that

$$\gamma_{4,m} - \gamma_{3,m} + \gamma_{2,m} - \gamma_{1,m} = 0 \implies \Delta_m = 0 \,. \tag{5.65}$$

We therefore expect a slowdown of the dynamics close to the hyperplane $\gamma_{4,m} - \gamma_{3,m} + \gamma_{2,m} - \gamma_{1,m} = 0$. In Fig. 5.8 we show that, for the two targets $(h = 0.5, \delta = 0.5)$ and $(h = 0, \delta = 0)$, such a slowdown is actually seen.

5.4 Variational bounds for disordered Ising chains

So far we have considered translationally invariant cases. Indeed, translational invariance was heavily used in the derivation of the variational bounds. The natural question is: can we prove similar bounds in presence of disorder. Here we show

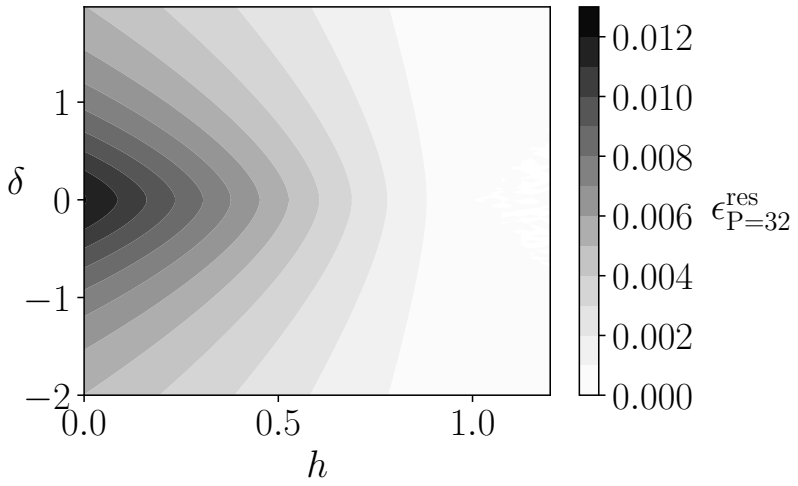

FIGURE 5.7: Digital ground state preparation for the XY model. Optimal residual energy attainable with depth P = 32 as a function of the target state parameters. Darker regions are harder to reach.

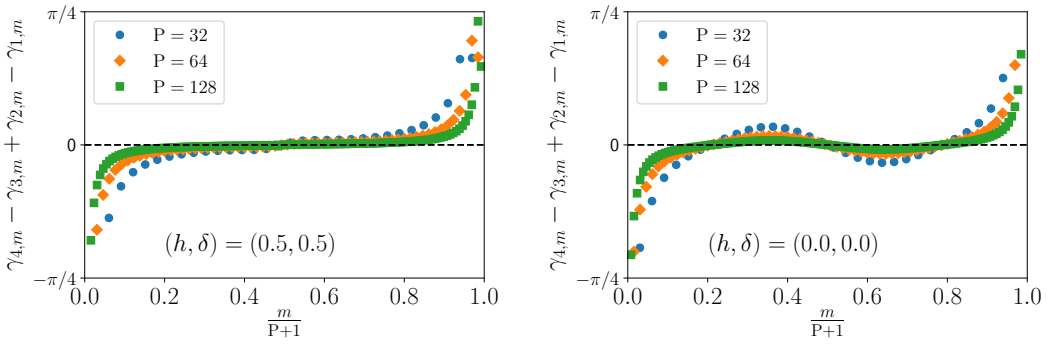

FIGURE 5.8: Optimal dQA trajectory when targeting the point $(h = 0.5, \delta = 0.5)$ and $(h = 0, \delta = 0)$ in the XY phase diagram with the four driving Hamiltonians in Eq. (5.62). We plot the parameter $\phi = \gamma_{4,m} - \gamma_{3,m} + \gamma_{2,m} - \gamma_{1,m}$. The critical lines $|h| = 1$ in the XY phase digram correspond to the lines $\phi = n\frac{\pi}{2}$ (with integer n) in the effective Hamiltonian driving the digital dynamics. The circuits have $\text{P} = 32, 64, 128$ and the length of the chain is $N = 1024$.

that it is possible to derive similar bounds on the residual energy for disordered systems, as long as the Hamiltonian is local. However these bounds are in general not tight. We now show how to derive this more general bounds in the specific case of the transverse field Ising ring with random nearest-neighbor couplings.

We consider the digitized evolution obtained alternating

$$\widehat{H}_1 = -\sum_{j=1}^{N} \hat{\sigma}_j^x \tag{5.66}$$

$$\widehat{H}_2 = \sum_{j=1}^{N} J_j \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z \,, \tag{5.67}$$

where $J_j \in (0,1)$. As in the transnational invariant case, we start from $\mathbf{g}_{\text{in}} = (1,0)$ and want to target the ground state of $\widehat{H}_2$, which is a symmetric superposition of the two classical Néel states

$$|\psi_{\text{gs}}\rangle = \frac{|\uparrow\downarrow\uparrow\cdots\uparrow\downarrow\rangle + |\downarrow\uparrow\downarrow\cdots\downarrow\uparrow\rangle}{\sqrt{2}} \,, \tag{5.68}$$

hence $\mathbf{g}_{\text{T}} = (0,1)$. It is convenient to express the expectation value of the energy as an average over N reduced sub-chains of length $L_{\text{R}} = 2P+2$:

$$\begin{aligned} E(\boldsymbol{\gamma};\mathbf{g}_{\text{T}}) &= \sum_{j=1}^{N} J_j \langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\sigma}_j^z\hat{\sigma}_{j+1}^z|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle \\ &= \frac{1}{L_{\text{R}}-1}\sum_{j_s=1}^{N}\sum_{j=j_s}^{j_s+L_{\text{R}}-2} J_j \langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\sigma}_j^z\hat{\sigma}_{j+1}^z|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle \\ &= \frac{1}{L_{\text{R}}-1}\sum_{j_s=1}^{N}\langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\mathcal{H}}^{(j_s)}|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle \\ &= \frac{1}{L_{\text{R}}-1}\sum_{j_s=1}^{N} E^{(j_s)}(\boldsymbol{\gamma},\mathbf{g}_{\text{T}}) \,, \end{aligned} \tag{5.69}$$

where j_s is the first site of each sub-chain and we have defined the reduced Hamiltonian $\hat{\mathcal{H}}^{(j_s)}$ and its expectation value on the final state $E^{(j_s)}(\boldsymbol{\gamma};\mathbf{g}_{\text{T}})$ to be

$$\hat{\mathcal{H}}^{(j_s)} = \sum_{j=j_s}^{j_s+L_{\text{R}}-2} J_j \hat{\sigma}_j^z\hat{\sigma}_{j+1}^z \tag{5.70}$$

$$E^{(j_s)}(\boldsymbol{\gamma},\mathbf{g}_{\text{T}}) = \langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\mathcal{H}}^{(j_s)}|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle \,. \tag{5.71}$$

We now focus on giving a lower bound for $E^{(j_s)}(\boldsymbol{\gamma},\mathbf{g}_{\text{T}})$. By using this bound in Eq. (5.69) we will get the desired lower bound for $E(\boldsymbol{\gamma};\mathbf{g}_{\text{T}})$.

We start by proving a trivial property of the state $|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle$, namely that

$$|j-j'|_{\text{ring}} \geq L_{\text{R}}-1 \implies \langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\sigma}_j^z\hat{\sigma}_{j'}^z|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle = 0 \,, \tag{5.72}$$

where $|j-j'|_{\text{ring}} = \min(|j-j'|, N-|j-j'|)$. To do that, we write the average in the Heisenberg's representation:

$$\langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\sigma}^z_j\hat{\sigma}^z_{j'}|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle = \langle\psi_0|\hat{\sigma}^z_j(\boldsymbol{\gamma})\hat{\sigma}^z_{j'}(\boldsymbol{\gamma})|\psi_0\rangle \ , \tag{5.73}$$

where

$$\hat{\sigma}^z_j(\boldsymbol{\gamma}) = \widehat{U}^\dagger(\boldsymbol{\gamma}_1)\cdots\widehat{U}^\dagger(\boldsymbol{\gamma}_{\text{P}})\hat{\sigma}^z_j\widehat{U}(\boldsymbol{\gamma}_{\text{P}})\cdots\widehat{U}(\boldsymbol{\gamma}_1) \ . \tag{5.74}$$

By simplifying the terms that commute, we observe that $\hat{\sigma}^z_j(\boldsymbol{\gamma})$ only involves the following $L_{\text{R}}-1$ spins

$$\left\{j-\left(\frac{L_{\text{R}}}{2}-1\right),\cdots,j+\left(\frac{L_{\text{R}}}{2}-1\right)\right\} \ .$$

In particular if $|j-j'|_{\text{ring}} \geq L_{\text{R}}-1$, $\hat{\sigma}^z_j(\boldsymbol{\gamma})$ and $\hat{\sigma}^z_{j'}(\boldsymbol{\gamma})$ involve two disjoint sets of spins. Recalling that $|\psi_0\rangle$ is a product state, we conclude that if $|j-j'|_{\text{ring}} \geq L_{\text{R}}-1$:

$$\begin{aligned}\langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\sigma}^z_j\hat{\sigma}^z_{j'}|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle &= \langle\psi_0|\hat{\sigma}^z_j(\boldsymbol{\gamma})\hat{\sigma}^z_{j'}(\boldsymbol{\gamma})|\psi_0\rangle \\ &= \langle\psi_0|\hat{\sigma}^z_j(\boldsymbol{\gamma})|\psi_0\rangle\langle\psi_0|\hat{\sigma}^z_{j'}(\boldsymbol{\gamma})|\psi_0\rangle \\ &= \langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\sigma}^z_j|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle\langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\sigma}^z_{j'}|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle = 0 \ ,\end{aligned} \tag{5.75}$$

where the final equality is a consequence of spin-inversion symmetry of the problem. This proves Eq. (5.72). In particular, it follows that, for the reduced problem, the first and last spins are uncorrelated

$$\langle\psi_{\text{P}}(\boldsymbol{\gamma})|\hat{\sigma}^z_{j_s}\hat{\sigma}^z_{j_s+L_{\text{R}}-1}|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle = 0 \ . \tag{5.76}$$

As a consequence of Eq. (5.76), the first and last spins of the reduced chain cannot be coherently anti-aligned and $|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle$ *must* have a non-vanishing overlap with the excited states of the reduced problem Hamiltonian $\hat{\mathcal{H}}^{(j_s)}$. To obtain a quantitative bound from this observation, we define $p^{(j_s)}_{\text{gs}}$ to be the probability that $|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle$ is found in one of the various ground states of $\hat{\mathcal{H}}^{(j_s)}$. (We note that degeneracy of the ground state is $2^{N-L_{\text{R}}+1}$ because a flip of a spin which lies outside of the reduced chain does not create excitations.) We can bound $p^{(j_s)}_{\text{gs}}$ with the probability $p_{\text{c}}\left(\sigma^z_{j_s}\sigma^z_{j_s+L_{\text{R}}-1}=-1\right)$ that the spins on sites j_s and $j_s+L_{\text{R}}-1$ are anti-aligned when the state of the system is $|\psi_{\text{P}}(\boldsymbol{\gamma})\rangle$. Indeed, since in all the

ground states this spins are anti-aligned, we can write

$$\begin{aligned} p_{\rm gs}^{(j_s)} &\leq p_{\rm c}\left(\sigma_{j_s}^z\sigma_{j_s+L_{\rm R}-1}^z=-1\right) \\ &= \frac{1}{2}\langle\psi_{\rm P}(\boldsymbol{\gamma})|1-\hat{\sigma}_{j_s}^z\hat{\sigma}_{j_s+L_{\rm R}-1}^z|\psi_{\rm P}(\boldsymbol{\gamma})\rangle=\frac{1}{2}\,. \end{aligned} \tag{5.77}$$

Some simple algebra then implies that

$$\begin{aligned} E^{(j_s)}(\boldsymbol{\gamma},\mathbf{g}_{\rm T}) &= \langle\psi_{\rm P}(\boldsymbol{\gamma})|\hat{\mathcal{H}}^{(j_s)}|\psi_{\rm P}(\boldsymbol{\gamma})\rangle \\ &\geq E_{\rm gs}^{(j_s)}+\Delta^{(j_s)}(1-p_{\rm gs}^{(j_s)}) \\ &\geq E_{\rm gs}^{(j_s)}+\frac{1}{2}\Delta^{(j_s)} \end{aligned} \tag{5.78}$$

where $\Delta^{(j_s)}>0$ is the energy gap between the ground states and the first excited states of $\hat{\mathcal{H}}^{(j_s)}$. As long as $N\geq L_{\rm R}$, we can insert this inequality into the expression for the energy of the whole system and get a variational lower bound for the total energy

$$E(\boldsymbol{\gamma};\mathbf{g}_{\rm T})\geq E_{\rm gs}(\mathbf{g}_{\rm T})+\frac{1}{2(L_{\rm R}-1)}\sum_{j_s=1}^{N}\Delta^{(j_s)}\,, \tag{5.79}$$

and the residual energy

$$\epsilon_{\rm P}^{\rm res}(\boldsymbol{\gamma},\mathbf{g}_{\rm T})\geq\frac{1}{L_{\rm R}-1}\frac{\frac{1}{2N}\sum_{j_s=1}^{N}\Delta^{(j_s)}}{e_{\rm max}(\mathbf{g}_{\rm T})-e_{\rm gs}(\mathbf{g}_{\rm T})}\,. \tag{5.80}$$

For the random Ising chain the numerator and denominator appearing in Eq. (5.80) takes the form

$$\frac{1}{2N}\sum_{j_s=1}^{N}\Delta^{(j_s)} = \frac{1}{N}\sum_{j_s=1}^{N}\min_{j\in[j_s,j_s+L_{\rm R}-2]}J_j \tag{5.81}$$

$$e_{\rm gs}(\mathbf{g}_{\rm T})-e_{\rm max}(\mathbf{g}_{\rm T}) = \frac{2}{N}\sum_{j=1}^{N}J_j\,. \tag{5.82}$$

Having the expression for the bound, we can specify it to different distributions of couplings. Here we discuss the case where the couplings are independent random variables, uniformly distributed in the interval $J_j\in[J_{\rm min},J_{\rm max}]$ with $J_{\rm max}\geq J_{\rm min}\geq 0$. We consider the infinite system limit $N\to\infty$ (or equivalently $N\gg(2P+2)$). Using the properties of the uniform distribution it is straightforward to derive the

following relations

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{2N}\sum_{j=1}^{N}\Delta^{(j_s)} &= \lim_{N\to\infty}\frac{1}{N}\sum_{j_s=1}^{N}\min_{j\in[j_s,j_s+L_{\mathrm{R}}-2]}J_j \\ &= \langle\min_{j\in[1,L_{\mathrm{R}}-1]}J_j\rangle_D \\ &= J_{\min}+\frac{J_{\max}-J_{\min}}{L_{\mathrm{R}}} \end{aligned} \tag{5.83}$$

$$\lim_{N\to\infty}\frac{2}{N}\sum_{j=1}^{N}J_j = J_{\max}+J_{\min}\,. \tag{5.84}$$

The bound then assumes the form

$$\begin{aligned}\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\mathbf{g}_{\mathrm{T}}) &\geq \frac{J_{\min}+\frac{J_{\max}-J_{\min}}{2\mathrm{P}+2}}{J_{\min}+J_{\max}}\frac{1}{2\mathrm{P}+1} = \\ &= \frac{a}{2\mathrm{P}+1}+\frac{b}{(2\mathrm{P}+1)(2\mathrm{P}+2)}\,, \end{aligned} \tag{5.85}$$

where

$$a=\frac{J_{\min}}{J_{\min}+J_{\max}}\ \text{ and }\ b=\frac{J_{\max}-J_{\min}}{J_{\min}+J_{\max}}\,. \tag{5.86}$$

For the homogeneous chain $J_{\min}=J_{\max}=1$ and we get:

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\mathbf{g}_{\mathrm{T}})\geq\frac{1}{2}\frac{1}{2\mathrm{P}+1}=\frac{1}{4\mathrm{P}+2}\,, \tag{5.87}$$

where we observe that the bound is less tight than that derived using translational invariance. We also note that the variational bound for the residual energy of a random chain scales as P^{-2} if $J_{\min}=0$ (since $a=0$).

This intriguing fact can be perhaps better understood from the limiting case where $J_j=0$ with unit probability. In this case the residual energy is strictly null for all schedules. On the other hand the density of defects ρ_{def}, which for an anti-ferromagnetic chain is defined by

$$\rho_{\mathrm{def}}=\frac{1}{N}\sum_{j=1}^{N}\langle\psi_{\mathrm{P}}(\boldsymbol{\gamma})|\frac{1-\hat{\sigma}_j^z\hat{\sigma}_{j+1}^z}{2}|\psi_{\mathrm{P}}(\boldsymbol{\gamma})\rangle \tag{5.88}$$

will always satisfy the bound in Eq. (5.87). This result can be obtained with a computation similar to the one presented for the residual energy.

5.5 Continuous-time dynamics and Lieb-Robinson bound

Let us return here to the issue of a continuous-time dynamics. We recall that when continuous-time ground state preparation protocols are considered, one must minimize the functional $E_\tau[\mathbf{g}(t)]$ given in Eq. (5.10). Our goal is to generalize to this case the residual energy bound for digital preparation obtained in section 5.2.1.

Since Eq 5.17 only relies on translational invariance, it also holds for a continuous-time protocol and takes the form

$$E_\tau[\mathbf{g}(t)] = N \sum_{\mu=1}^{\mathrm{n_c}} g_\mu \langle \psi(\tau)|\widehat{O}_\mu(\mathbf{r}_0)|\psi(\tau)\rangle \ . \tag{5.89}$$

The residual energy bound found for the digital dynamics emerged by combining Eq 5.89 with the existence of an exact light-cone (operator spreading) picture. In the continuous-time dynamics the operator spreading follows an approximate light cone [Lieb and Robinson, 1972]. As we will show in the next section, this is sufficient to derive a variational bound for $E_\tau[\mathbf{g}(t)]$.

5.5.1 Lieb-Robinson and quasi-locality of the dynamics

In Lieb and Robinson [1972] Lieb and Robinson formalized the intuitive notion that in a system with bounded local interactions there is some limit, v, to the speed with which correlations can propagate. In recent years, various rigorous upper bounds to v have been proved, which usually go under the name of "Lieb-Robinson bounds".

In particular Refs. [Nachtergaele and Sims, 2009; Barthel and Kliesch, 2012; Kliesch *et al.*, 2014] showed that the dynamics induced by a system with a short-range time-dependent Hamiltonian is quasi-local. This means that (when $\mathbf{g}(t)$ is bounded), up to an exponentially small error, expectation values of local operators can be evaluated using a Hamiltonian defined on a reduced system whose size L_{R} scales linearly in time $L_{\mathrm{R}} > vt$. In the specific case considered here, using Eq. (11) and Eq. (14) from Barthel and Kliesch [2012] we get following inequality

$$\begin{aligned} &\left|\langle \psi(\tau)|\widehat{O}_\mu(\mathbf{r}_0)|\psi(\tau)\rangle - \langle \widetilde{\psi}(\tau)|\widehat{O}_\mu(\mathbf{r}_0)|\widetilde{\psi}(\tau)\right| \\ &\qquad \leq \frac{C}{2}||\widehat{O}_\mu(\mathbf{r}_0)|| \left(\frac{L_{\mathrm{R}}}{2\ell}\right)^{d-1} \mathrm{e}^{v\tau - \frac{L_{\mathrm{R}}}{2\ell}} \ , \end{aligned} \tag{5.90}$$

where the constant C depends on the range of the interaction, denoted by ℓ, and on the geometry of the lattice (we have here assumed a d-dimensional hypercubic lattice), but they do not depend on the system size nor on time. Here $||\cdot||$ is the standard operator norm (modulus of largest eigenvalue).

Since the reduced system is insensitive to the boundary conditions, we may choose arbitrary boundary condition on the reduced Hamiltonian. In particular the choice of a set of boundary conditions compatible with translational invariance, as in Eqs. (5.26), (5.27) and (5.32), results, by using a triangle inequality, in the relation

$$\begin{aligned}\left|\langle\psi(\tau)|\widehat{H}(\mathbf{g})|\psi(\tau)\rangle - \langle\widetilde{\psi}(\tau)|\hat{\mathcal{H}}^{(s_x s_y)}(\mathbf{g})|\widetilde{\psi}(\tau)\right| \\ \leq C\sum_{\mu=1}^{n_c}||\widehat{O}_\mu(\mathbf{r}_0)||\left(\frac{L_{\rm R}}{2\ell}\right)^{d-1}\mathrm{e}^{v\tau-\frac{L_{\rm R}}{2\ell}}\,.\end{aligned} \tag{5.91}$$

By applying the variational principle we get a bound on the residual energy:

$$\begin{aligned}e_\tau[\mathbf{g}(t),\mathbf{g}_{\rm T}] &\geq \max_{s_x,s_y}\frac{\mathcal{E}_{\rm gs}^{(s_x s_y)}(\mathbf{g}_{\rm T})}{N_{\rm R}} + \widetilde{C}L_{\rm R}^{d-1}e^{v\tau-\frac{L_{\rm R}}{2\ell}} \\ &\stackrel{\rm def}{=} \varepsilon_{\rm R}(\mathbf{g}_{\rm T}) + \widetilde{C}L_{\rm R}^{d-1}e^{v\tau-\frac{L_{\rm R}}{2\ell}}\,,\end{aligned} \tag{5.92}$$

where $\widetilde{C}$ collects all numerical prefactors. We conclude that, as long as $L > L_{\rm R}$:

$$\epsilon_\tau^{\rm res}[\mathbf{g}(t),\mathbf{g}_{\rm T}] \geq \epsilon_{\rm P}^{\rm var}(\mathbf{g}_{\rm T}) + \frac{\widetilde{C}L_{\rm R}^{d-1}e^{v\tau-\frac{L_{\rm R}}{2\ell}}}{e_{\rm max}(\mathbf{g}_{\rm T}) - e_{\rm gs}(\mathbf{g}_{\rm T})}\,, \tag{5.93}$$

where

$$\epsilon_{\rm P}^{\rm var}(\mathbf{g}_{\rm T}) = \frac{\varepsilon_{\rm R}(\mathbf{g}_{\rm T}) - e_{\rm gs}(\mathbf{g}_{\rm T})}{e_{\rm max}(\mathbf{g}_{\rm T}) - e_{\rm gs}(\mathbf{g}_{\rm T})}\,. \tag{5.94}$$

Notice that here $L_{\rm R}$ is a constant that enters implicitly in $\epsilon_{\rm P}^{\rm var}(\mathbf{g}_{\rm T})$ and which should be chosen in such a way that the exponential term is small, to get a meaningful bound.

If for example we take $L_{\rm R} = 4\ell v\tau < L$ (which is always possible if $L \to \infty$), for the TFIM with target $h = 0$ we get (in any dimension)

$$\epsilon_\tau^{\rm res}[\mathbf{g}(t),\mathbf{g}_{\rm T}] \geq \frac{1}{4\ell v\tau} + \mathrm{Const}\,(v\tau)^{d-1}\,e^{-v\tau}\,. \tag{5.95}$$

This shows that the residual energy cannot scale better than τ^{-1} independently of the chosen schedule.

5.6 KZ scaling and adiabatic perturbation theory: a comparison

The considerations of the previous section call for a detailed comparison with previously known results on the Kibble-Zurek mechanism. Indeed, when a system is driven across a quantum critical point, defects n^{def} are created due to non-adiabatic transitions. According to the Kibble-Zurek (KZ) mechanism, for a system in d dimensions, the density of defects created when crossing a quantum critical point at a finite rate τ^{-1} scales as [Polkovnikov, 2005; De Grandi and Polkovnikov, 2010]

$$n_{\tau}^{\text{def,KZ}} \sim \tau^{-\frac{\nu d}{1+z\nu}} , \tag{5.96}$$

where ν snd z are the universal critical exponents that characterize the quantum critical point. For the TFIM $z = 1$ while ν depends on the d:

$$\begin{cases} z = 1 \quad \nu = 1, & \text{for } d = 1 \\ z = 1 \quad \nu \approx 0.63 & \text{for } d = 2 \\ z = 1 \quad \nu = \frac{1}{2} & \text{for } d \geq 3 \end{cases} . \tag{5.97}$$

Hence, the KZ mechanism for the TFIM would predict:

$$n_{\tau}^{\text{def,KZ}} = \begin{cases} \sim \tau^{-\frac{1}{2}}, & \text{for } d = 1 \\ \sim \tau^{-0.773} & \text{for } d = 2 \\ \sim \tau^{-\frac{d}{3}} & \text{for } d \geq 3 \end{cases} . \tag{5.98}$$

In Refs. [Polkovnikov, 2005; De Grandi and Polkovnikov, 2010] a derivation of the KZ scaling using only Adiabatic Perturbation Theory (APT) and the usual scaling relations which hold in proximity of critical points is given. Since the adiabatic perturbation theory is correct up to an order $\mathcal{O}(\tau^{-2})$, the breaking of the KZ scaling was predicted when such τ^{-2} corrections become dominant. In

particular when applied to the TFIM, this theory would predict

$$n_\tau^{\text{def,APT}} = \begin{cases} \sim \tau^{-\frac{1}{2}}, & \text{for } d = 1 \\ \sim \tau^{-0.773} & \text{for } d = 2 \\ \sim \tau^{-1} & \text{for } d = 3 \\ \sim \tau^{-\frac{4}{3}} & \text{for } d = 4 \\ \sim \tau^{-\frac{5}{3}} & \text{for } d = 5 \\ \sim \tau^{-2} & \text{for } d \geq 6 \end{cases} . \tag{5.99}$$

The scaling of the residual energy additionally depends on the dispersion relation of the target Hamiltonian. However, it can be easily computed for two specific cases: the critical point h_c and the classical state with $h = 0$. If the target Hamiltonian is the critical point, we can use the dispersion relation $\epsilon_{\mathbf{k}} \sim |\mathbf{k}|^z$, which leads to

$$\epsilon_\tau^{\text{def,APT}}[h_c] \sim \tau^{-\min(\frac{\nu(d+z)}{1+z\nu},2)} = \begin{cases} \sim \tau^{-1}, & \text{for } d = 1 \\ \sim \tau^{-1.159} & \text{for } d = 2 \\ \sim \tau^{-\frac{4}{3}} & \text{for } d = 3 \\ \sim \tau^{-\frac{5}{3}} & \text{for } d = 4 \\ \sim \tau^{-2} & \text{for } d \geq 5 \end{cases} \tag{5.100}$$

We observe that the bound that we have found at the critical point, in $d = 1$, has indeed a larger power τ^{-2}, hence definitely better than the linear schedule KZ result. For higher d we are unfortunately unable to carry out detailed calculations.

The other case that can be easily addressed is when the target has $h = 0$. In this case, one can argue that the final residual energy is directly proportional to the density of excitations at the critical point, hence we expect:

$$\epsilon_\tau^{\text{def,APT}}[h = 0] \sim \tau^{-\min(\frac{\nu d}{1+z\nu},2)} = \begin{cases} \sim \tau^{-\frac{1}{2}}, & \text{for } d = 1 \\ \sim \tau^{-0.773} & \text{for } d = 2 \\ \sim \tau^{-1} & \text{for } d = 3 \\ \sim \tau^{-\frac{4}{3}} & \text{for } d = 4 \\ \sim \tau^{-\frac{5}{3}} & \text{for } d = 5 \\ \sim \tau^{-2} & \text{for } d \geq 6 \end{cases} \tag{5.101}$$

The bound we derived in this work suggests that when the target is at $h = 0$ the residual energy cannot scale better than $\sim \tau^{-1}$. This would be in disagreement with the APT results — which are in principle tailored to exploring properties close to the critical point — for $d > 3$.

Chapter 6

Conclusions

This Thesis dealt with some of the issues concerning Quantum Annealing and other digital quantum state preparation algorithms, notably the Quantum Approximate Optimization Algorithm (QAOA).

We started, in Chap. 2, by re-examining the relationship between the Simulated QA (SQA), implemented with a Path-Integral Monte Carlo dynamics, and the actual Schrödinger QA dynamics. We did that in the specific case of the disordered quantum Ising chain, where analytical tools allowed us to simulate the Schrödinger dynamics of large systems. The results show that, although the correct quantum mechanical treatment does require the limit $P \to \infty$, this by itself does not guarantee that the resulting SQA dynamics is *physical*. In particular, the typical "slow-down" that SQA with $P \to \infty$ tends to show for large annealing times — which we have verified to be due to disorder, even in absence of frustration and true complexity — does not necessarily imply that one should expect no quantum speed-up from the actual QA dynamics.

Next, we moved to a series of studies that provide a link between traditional QA and alternative hybrid quantum-classical approaches, such as QAOA, with emphasis on optimal quantum control and schedule optimization. Our contribution to the field was twofold.

As a first contribution, we developed a technique to establish a variational bound on the performance of QAOA for translationally invariant spin problems on regular periodic graphs. We did this by exploiting the intrinsic flexibility in the boundary conditions of a reduced spin problem to prove rigorous bounds on the relative error ϵ_P^{res} of the QAOA optimal solution on a circuit of depth P.

As a second contribution, we establish a link between optimal quantum control and the adiabatic dynamics behind *digitized*-QA. Indeed, using the QAOA

framework, we explicitly showed that one can iteratively single-out a *smooth regular* solution which can be regarded as an *optimal digitized*-QA schedule, satisfying all the expected requirements for adiabaticity in a digital context, without any need for spectral information on the Hamiltonian. Moreover, the construction of the *optimal digitized*-QA schedule turned out to be computationally less expensive than searching for unstructured QAOA solutions.

To show the potential applications of the techniques developed — without having to wait for suitable quantum devices, i.e., exploiting currently available numerical techniques, such as DMRG Density Matrix Renormalization Group (DMRG) [Schollwöck, 2011] or other Tensor Network algorithms — consider the problem of preparing approximate ground state wave-functions for the Spin-1/2 XZZ Heisenberg chain with a magnetic field along the x-direction:

$$\widehat{H} = \sum_{j=1}^{N} \left[J(\hat{S}_j^z \hat{S}_{j+1}^z + \hat{S}_j^y \hat{S}_{j+1}^y) + J_x \hat{S}_j^x \hat{S}_{j+1}^x \right] - h \sum_j \hat{S}_j^x \ , \tag{6.1}$$

where $\hat{S}_j^{x,y,z} = \hat{\sigma}_j^{x,y,z}/2$ are spin-1/2 operators. [1] The phase diagram of this model, shown in Fig. 6.1, is rather rich. [2] It includes (for $J_x = 0$) the free-fermion line associated to $\delta = 0$ in the XY phase diagram we discussed in Chapter 5: hence, the origin ($h = 0, J_x = 0$) is a point we have already shown how to prepare. The other axis (for $h = 0$) is the standard Heisenberg chain, which includes the critical (Luttinger) line and the Nèel anti-ferromagnet (along the x-direction) for large $J_x/J > 1$. This makes the model an ideal testing ground to benchmark ground state preparation algorithms beyond the free-fermion framework we have explored so far. Here one might develop bounds and calculate approximate wave-functions by using the *optimal digitized*-QA approach outlined, with suitable generalizations (possibly by DMRG) to account for the $\hat{S}_j^x \hat{S}_{j+1}^x$ terms which, when tackled with a Jordan-Wigner transformation, leads to interactions between fermions.

Another interesting issue has to do with the role of disorder. We have verified, and will report elsewhere, that the perfect degeneracy of the optimal QAOA solutions found in the translationally invariant quantum Ising chain is broken in the presence of disorder: the variational energy landscape becomes extremely rugged,

[1] Notice that, as written, the anisotropy is in the $\hat{S}_j^x \hat{S}_{j+1}^x$ term, as opposed to the more standard $\hat{S}_j^z \hat{S}_{j+1}^z$ term. The reason for that is that we insist in using the standard "transverse field" $\hat{S}_j^x$ of QA.

[2] The phase diagram is based on *Bethe Ansatz* results provided in Yang and Yang [1966]. Rakov and Weyrauch [2019] used a Tensor Renormalization Group approach to study the low-energy properties of the model.

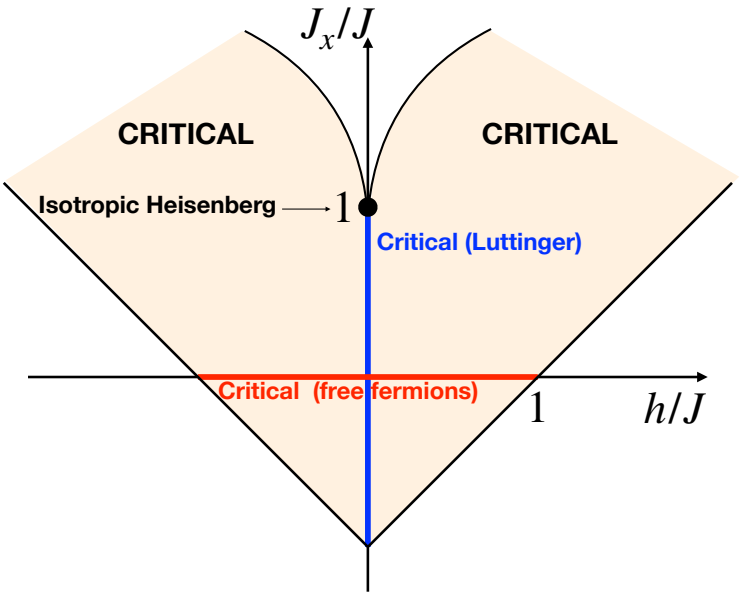

FIGURE 6.1: Phase diagram of the Spin-1/2 XZZ Heisenberg chain in a magnetic field, see Eq. 6.1, in the plane (h, J_x). On the $J_x = 0$ line the system is described by free fermions which are obtained by setting $\delta = 0$ in the XY model studied in Chapter 5. The colored region corresponds to a critical (i.e., gapless) phase. Notice also the presence of Anti-ferromagnetic (AF) Néel long range ordered (LRO) phase for large $J_x/J > 1$.

and the search for the global optimal solution turns to be a computationally hard problem. Further scrutiny is needed to investigate the quality of the *optimal digitized*-QA solution in a situation in which a large number of non-degenerate minima is present. The application of Machine Learning ideas Bukov *et al.* [2018]; Wecker *et al.* [2016]; Crooks [2018] to such complex minimization problems appears to be a fascinating perspective.

Appendix A

Diagonalization of the one-dimensional XY model

The Jordan-Wigner (JW) transformation for the XY chain is rather standard. For QA and QAOA in a chain with PBC, see for instance Wang *et al.* [2018] and Dziarmaga [2005]. In this Thesis, however, the boundary conditions play a crucial role. In particular, various results were obtained by considering a reduced spin chain with ABC rather than PBC. In the following, we present a unified derivation, valid for both PBC and ABC.

To set the notation used in this Appendix, we consider an XY chain of length $N_{\rm R}$, with PBC (+) and ABC (−). The system Hamiltonian is then a linear combination of the following three driving Hamiltonians:

$$\hat{\mathcal{H}}_x^{(\pm)} = -\sum_{j=1}^{N_{\rm R}} \hat{\sigma}_j^x \tag{A.1}$$

$$\hat{\mathcal{H}}_{yy}^{(\pm)} = \sum_{j=1}^{N_{\rm R}-1} \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y \pm \hat{\sigma}_{N_{\rm R}}^y \hat{\sigma}_1^y \tag{A.2}$$

$$\hat{\mathcal{H}}_{zz}^{(\pm)} = \sum_{j=1}^{N_{\rm R}-1} \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z \pm \hat{\sigma}_{N_{\rm R}}^z \hat{\sigma}_1^z \, . \tag{A.3}$$

The specific case of the quantum Ising chain, considered in Chapter 3 and Chapter 4, is obtained discarding $\hat{\mathcal{H}}_{yy}^{(\pm)}$.

A.1 Diagonalization of the Hamiltonian

The global parity $\widehat{\mathcal{P}} = \prod_{n=1}^{N_R} \hat{\sigma}_n^x$ is a conserved quantity for all the Hamiltonian terms. Therefore, since the initial state has an even parity $\widehat{\mathcal{P}}|\psi_0\rangle = |\psi_0\rangle$, we can restrict our analysis to the even parity subspace. When restricted to the even parity sector, a Jordan-Wigner (JW) transformation Jordan and Wigner [1928]

$$\hat{\sigma}_j^x = 1 - 2\hat{c}_j^\dagger \hat{c}_j \tag{A.4}$$

$$\hat{\sigma}_j^y = -i(\hat{c}_j - \hat{c}_j^\dagger) \exp\left(-i\pi \sum_{l=1}^{j-1} \hat{c}_l^\dagger \hat{c}_l\right) \tag{A.5}$$

$$\hat{\sigma}_j^z = -(\hat{c}_j + \hat{c}_j^\dagger) \exp\left(-i\pi \sum_{l=1}^{j-1} \hat{c}_l^\dagger \hat{c}_l\right), \tag{A.6}$$

maps the spin system to free fermions on a lattice, where $\hat{c}_j^\dagger$ and $\hat{c}_j$, respectively, create and annihilate a spinless fermion at site j. After the JW transformation, the various Hamiltonians take the form

$$\hat{\mathcal{H}}_x^{(\pm)} = \sum_{n=1}^{N_R} (\hat{c}_j^\dagger \hat{c}_j - \hat{c}_j \hat{c}_j^\dagger) \tag{A.7}$$

$$\hat{\mathcal{H}}_{yy}^{(\pm)} = \sum_{n=1}^{N_R-1} (\hat{c}_j^\dagger \hat{c}_{j+1} - \hat{c}_j^\dagger \hat{c}_{j+1}^\dagger + \text{H.c.}) \mp (\hat{c}_{N_R}^\dagger \hat{c}_1 - \hat{c}_{N_R}^\dagger \hat{c}_1^\dagger + \text{H.c.}) \tag{A.8}$$

$$\hat{\mathcal{H}}_{zz}^{(\pm)} = \sum_{n=1}^{N_R-1} (\hat{c}_j^\dagger \hat{c}_{j+1} + \hat{c}_j^\dagger \hat{c}_{j+1}^\dagger + \text{H.c.}) \mp (\hat{c}_{N_R}^\dagger \hat{c}_1 + \hat{c}_{N_R}^\dagger \hat{c}_1^\dagger + \text{H.c.}), \tag{A.9}$$

where H.c. stands for Hermitian conjugate. We observe that PBC for the spins are mapped into ABC for the fermions and vice-versa The parity operator translates into:

$$\widehat{\mathcal{P}} = \exp\left(-i\pi \sum_{l=1}^{N_R} \hat{c}_l^\dagger \hat{c}_l\right) = 1 \quad \text{(valid in the sector of even parity)}. \tag{A.10}$$

A Fourier transform can then be used to decompose the system in a set of decoupled two level systems. This is done by introducing a set of wave-vectors $\tilde{\mathcal{K}}^{(\pm)}$ that, to be consistent with the BC, must be taken to be

$$\text{spin} - \text{PBC}: \quad \tilde{\mathcal{K}}^{(+)} = \left\{ \pm\pi \frac{2n-1}{N_R} \text{ for } n = 1, 2, \cdots, \frac{N_R}{2} \right\} \tag{A.11}$$

$$\text{spin} - \text{ABC}: \quad \tilde{\mathcal{K}}^{(-)} = \left\{ \pm 2\pi \frac{n}{N_R} \text{ for } n = 1, 2, \cdots, \frac{N_R}{2} - 1 \right\} \cup \left\{ 0, \pi \right\} \tag{A.12}$$

and substituting

$$\hat{c}_j^\dagger = \frac{\mathrm{e}^{i\pi/4}}{\sqrt{N_\mathrm{R}}} \sum_k^{\tilde{\mathcal{K}}^{(\pm)}} e^{+ikj} \hat{c}_k^\dagger \,, \tag{A.13}$$

where $\hat{c}_k^\dagger$ creates a fermion of wave-vector k.

In terms of the Fourier modes, using that $\mathrm{e}^{ikN_\mathrm{R}} = \mp 1$ for $k \in \tilde{\mathcal{K}}^{(\pm)}$, the Hamiltonians read

$$\hat{\mathcal{H}}_x^{(\pm)} = \sum_k^{\tilde{\mathcal{K}}^{(\pm)}} (\hat{c}_k^\dagger \hat{c}_k - \hat{c}_k \hat{c}_k^\dagger) \tag{A.14}$$

$$\hat{\mathcal{H}}_{yy}^{(\pm)} = \sum_k^{\tilde{\mathcal{K}}^{(\pm)}} 2\cos k\, \hat{c}_k^\dagger \hat{c}_k - \sum_k^{\tilde{\mathcal{K}}^{(\pm)}} \sin k\, (\hat{c}_k^\dagger \hat{c}_{-k}^\dagger + \mathrm{H.c.}) \tag{A.15}$$

$$\hat{\mathcal{H}}_{zz}^{(\pm)} = \sum_k^{\tilde{\mathcal{K}}^{(\pm)}} 2\cos k\, \hat{c}_k^\dagger \hat{c}_k + \sum_k^{\tilde{\mathcal{K}}^{(\pm)}} \sin k\, (\hat{c}_k^\dagger \hat{c}_{-k}^\dagger + \mathrm{H.c.})\,. \tag{A.16}$$

Eqs. (A.14), (A.15) and (A.16) already show that the system decomposes into pairs of modes with conjugate momenta k and $-k$. The main difference between PBC and ABC emerges at this level. In fact the exceptional modes $k = 0, \pi$, which appear only with spin-ABC, are self conjugate and do not couple to any other mode at all. A direct consequence of this is that, with ABC, the number operators associated with such modes are conserved quantities. In particular, since these modes are absent in the initial state $\hat{c}_0^\dagger \hat{c}_0 |\psi_0\rangle = \hat{c}_\pi^\dagger \hat{c}_\pi |\psi_0\rangle = 0$, we can confine ourselves to the subspace where the exceptional modes are absent.

For the spin-PBC case we notice that, due to the parity in the summed terms, we can write:

$$\sum_k^{\tilde{\mathcal{K}}^{(+)}} (\hat{c}_k^\dagger \hat{c}_k - \hat{c}_k \hat{c}_k^\dagger) = \sum_k^{\tilde{\mathcal{K}}^{(+)}} (\hat{c}_k^\dagger \hat{c}_k - \hat{c}_{-k} \hat{c}_{-k}^\dagger) = 2 \sum_{k>0}^{\tilde{\mathcal{K}}^{(+)}} (\hat{c}_k^\dagger \hat{c}_k - \hat{c}_{-k} \hat{c}_{-k}^\dagger) \tag{A.17}$$

where now the sum is restricted to positive $k > 0$. For the spin-ABC, however, the exceptional modes give an extra contribution -2 originating from the two terms $(2\hat{c}_{k=0,\pi}^\dagger \hat{c}_{k=0,\pi} - 1)$. All in all, we end up writing:

$$\hat{\mathcal{H}}_x^{(\pm)} = (-1 \pm 1) + 2 \sum_k^{\mathcal{K}^{(\pm)}} (\hat{c}_k^\dagger \hat{c}_k - \hat{c}_{-k} \hat{c}_{-k}^\dagger) \tag{A.18}$$

$$\tag{A.19}$$

where we have introduced the set of dynamically active (and positive) wave-vectors $\mathcal{K}^{(\pm)}$ given by

$$\text{PBC}: \ \mathcal{K}^{(+)} = \left\{ \pi \frac{2n-1}{N_{\text{R}}} \text{ for } n = 1, 2, \cdots, \frac{N_{\text{R}}}{2} \right\} \tag{A.20}$$

$$\text{ABC}: \ \mathcal{K}^{(-)} = \left\{ 2\pi \frac{n}{N_{\text{R}}} \text{ for } n = 1, 2, \cdots, \frac{N_{\text{R}}}{2} - 1 \right\} . \tag{A.21}$$

The anomalous term is naturally free from exceptional modes, as $\hat{c}^{\dagger}_{0,\pi}\hat{c}^{\dagger}_{0,-\pi} = 0$. The remaining terms being even in k can be written as:

$$\sum_{k}^{\tilde{\mathcal{K}}^{(\pm)}} \sin k \, (\hat{c}^{\dagger}_{k}\hat{c}^{\dagger}_{-k} + \text{H.c.}) = 2 \sum_{k}^{\mathcal{K}^{(\pm)}} \sin k \, (\hat{c}^{\dagger}_{k}\hat{c}^{\dagger}_{-k} + \text{H.c.}) \, . \tag{A.22}$$

Similar considerations apply to the $\cos k$ term, were we can write:

$$\begin{aligned} \sum_{k}^{\tilde{\mathcal{K}}^{(\pm)}} 2 \cos k \, \hat{c}^{\dagger}_{k}\hat{c}_{k} &= \sum_{k}^{\tilde{\mathcal{K}}^{(\pm)}} \cos k \, (\hat{c}^{\dagger}_{k}\hat{c}_{k} + \hat{c}^{\dagger}_{-k}\hat{c}_{-k}) \\ &= 2 \sum_{k}^{\mathcal{K}^{(\pm)}} \cos k \, (\hat{c}^{\dagger}_{k}\hat{c}_{k} - \hat{c}_{-k}\hat{c}^{\dagger}_{-k}) \, . \end{aligned} \tag{A.23}$$

Summarizing, the Hamiltonians read

$$\hat{\mathcal{H}}^{(\pm)}_{x} = (-1 \pm 1) + 2 \sum_{k}^{\mathcal{K}^{(\pm)}} (\hat{c}^{\dagger}_{k}\hat{c}_{k} - \hat{c}_{-k}\hat{c}^{\dagger}_{-k}) \tag{A.24}$$

$$\hat{\mathcal{H}}^{(\pm)}_{yy} = 2 \sum_{k}^{\mathcal{K}^{(\pm)}} \Big[\cos k \, (\hat{c}^{\dagger}_{k}\hat{c}_{k} - \hat{c}_{-k}\hat{c}^{\dagger}_{-k}) - \sin k \, (\hat{c}^{\dagger}_{k}\hat{c}^{\dagger}_{-k} + \hat{c}_{-k}\hat{c}_{k}) \Big] \tag{A.25}$$

$$\hat{\mathcal{H}}^{(\pm)}_{zz} = 2 \sum_{k}^{\mathcal{K}^{(\pm)}} \Big[\cos k \, (\hat{c}^{\dagger}_{k}\hat{c}_{k} - \hat{c}_{-k}\hat{c}^{\dagger}_{-k}) + \sin k \, (\hat{c}^{\dagger}_{k}\hat{c}^{\dagger}_{-k} + \hat{c}_{-k}\hat{c}_{k}) \Big] \, . \tag{A.26}$$

A further inspection of Eqs. (A.24)—(A.26) reveals that each pairs' parity operator $\widehat{\mathcal{P}}_{k} = \mathrm{e}^{i\pi(\hat{c}^{\dagger}_{k}\hat{c}_{k} + \hat{c}^{\dagger}_{-k}\hat{c}_{-k})}$ is also conserved. Again, since $\widehat{\mathcal{P}}_{k}|\psi_0\rangle = 1$ for all $k \in \mathcal{K}^{(\pm)}$, we can restrict our analysis to the subspace $\mathcal{P}_{k} = 1$ for all $k \in \mathcal{K}^{(\pm)}$. Finally, in this subspace, the system is equivalent to a collection of two-level systems. For instance, the identification

$$|\uparrow_{k}\rangle = |0\rangle \quad \text{and} \quad |\downarrow_{k}\rangle = \hat{c}^{\dagger}_{k}\hat{c}^{\dagger}_{-k}|0\rangle \tag{A.27}$$

maps the system into a collection of decoupled pseudo-spin-1/2. The number of independent pseudo-spins for PBC is $|\mathcal{K}^{(+)}| = N_{\mathrm{R}}/2$ while for ABC, due to the absence of the exceptional modes, the number of pseudo-spins is given by $|\mathcal{K}^{(-)}| = N_{\mathrm{R}}/2 - 1$. Introducing the pseudo-spin Pauli operators $\hat{\boldsymbol{\tau}}_k = (\hat{\tau}_x, \hat{\tau}_y, \hat{\tau}_z)^T$ with:

$$\hat{\tau}_k^x = \left(\hat{c}_k^\dagger \hat{c}_{-k}^\dagger + \hat{c}_{-k}\hat{c}_k\right) \tag{A.28}$$

$$\hat{\tau}_k^y = -i\left(\hat{c}_k^\dagger \hat{c}_{-k}^\dagger - \hat{c}_{-k}\hat{c}_k\right) \tag{A.29}$$

$$\hat{\tau}_k^z = -\left(\hat{c}_k^\dagger \hat{c}_k - \hat{c}_{-k}\hat{c}_{-k}^\dagger\right) , \tag{A.30}$$

the Hamiltonians read

$$\hat{\mathcal{H}}_x^{(\pm)} = (-1 \pm 1) + \sum_k^{\mathcal{K}^{(\pm)}} \hat{\mathcal{H}}_x^{(k)} \tag{A.31}$$

$$\hat{\mathcal{H}}_{yy}^{(\pm)} = \sum_k^{\mathcal{K}^{(\pm)}} \hat{\mathcal{H}}_{yy}^{(k)} \tag{A.32}$$

$$\hat{\mathcal{H}}_{zz}^{(\pm)} = \sum_k^{\mathcal{K}^{(\pm)}} \hat{\mathcal{H}}_{zz}^{(k)} , \tag{A.33}$$

where

$$\hat{\mathcal{H}}_x^{(k)} = -2\hat{\tau}_k^z = -2\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{\tau}}_k \tag{A.34}$$

$$\hat{\mathcal{H}}_{yy}^{(k)} = -2\sin k\, \hat{\tau}_k^x - 2\cos k\, \hat{\tau}_k^z = -2\hat{\boldsymbol{a}}_k \cdot \hat{\boldsymbol{\tau}}_k \tag{A.35}$$

$$\hat{\mathcal{H}}_{zz}^{(k)} = +2\sin k\, \hat{\tau}_k^x - 2\cos k\, \hat{\tau}_k^z = -2\hat{\boldsymbol{b}}_k \cdot \hat{\boldsymbol{\tau}}_k , \tag{A.36}$$

are the Hamiltonians for each independent pseudo-spin, and we defined the unit vectors $\hat{\boldsymbol{z}} = (0,0,1)^T$, $\hat{\boldsymbol{a}}_k = (\sin k, 0, \cos k)^T$ and $\hat{\boldsymbol{b}}_k = (-\sin k, 0, \cos k)^T$.

A.2 QAOA dynamics in quantum Ising chain

The pseudo-spin representation of the Hamiltonian is useful to discuss the digital dynamics induced by the QAOA Hamiltonians. As in Chapter 4, we focus on the quantum Ising chain and thus discard $\hat{\mathcal{H}}_{yy}^{(k)}$. We compute the residual energy associate with a variational state (see Eq. (A.37))

$$|\psi_{\mathrm{P}}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = \mathrm{e}^{-i\beta_{\mathrm{P}}\hat{H}_x}\mathrm{e}^{-i\gamma_{\mathrm{P}}\hat{H}_z} \cdots \mathrm{e}^{-i\beta_1\hat{H}_x}\mathrm{e}^{-i\gamma_1\hat{H}_z}|\psi_0\rangle , \tag{A.37}$$

in a quantum Ising chain of system of N spins. In particular we derive Eq. (4.20) and Eq. (4.21) from Chapter 4.

As discussed in Chapter 4 it is convenient to assume that N is even and carry out the computation in a reduced chain of $N_{\rm R} = \min(2{\rm P}+2, N)$. More specifically we should compute (see Eq. (4.17)):

$$\epsilon_{\rm P}^{\rm res}(\boldsymbol{\gamma},\boldsymbol{\beta}) \;=\; \langle\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})|\left(\frac{\hat{\mathcal{H}}_{zz}^{(\pm)}}{2(2{\rm P}+2)}+1\right)|\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle\;, \tag{A.38}$$

where we have used that $\hat{\mathcal{H}}_{z}^{(\pm)} = \hat{\mathcal{H}}_{zz}^{(\pm)}$. To simplify the notation, in the rest of this section, we omit the explicit indication of the boundary conditions used and the tilde to indicate reduced spin states. We will discuss the effect effect of the boundary

We introduce the unit vector describing the pseudo-spin magnetization

$$\boldsymbol{\tau}_k(\boldsymbol{\gamma},\boldsymbol{\beta}) = \langle\psi_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})|\hat{\boldsymbol{\tau}}_k|\psi_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle\;. \tag{A.39}$$

In the pseudo-spin representation, the initial state $|\psi_0\rangle$, being the ground state of $\hat{\mathcal{H}}_x$, corresponds to a state where all pseudo-spins are aligned along the $\hat{\boldsymbol{z}}$ axis. The initial pseudo-spin magnetization $\boldsymbol{\tau}_k(0)$ is therefore

$$\boldsymbol{\tau}_k(0) = \langle\psi_0|\hat{\boldsymbol{\tau}}_k|\psi_0\rangle = \hat{\boldsymbol{z}}\;. \tag{A.40}$$

Then, starting from the initial condition $\boldsymbol{\tau}_k(0) = \hat{\boldsymbol{z}}$, the QAOA Hamiltonians are used to perform a sequence of rotations on the pseudospins. The action of each Trotter step in the QAOA *Ansatz* is described by the identity

$$\mathrm{e}^{i\gamma_m\hat{\mathcal{H}}_{zz}}\mathrm{e}^{i\beta_m\hat{\mathcal{H}}_x}\hat{\boldsymbol{\tau}}_k\mathrm{e}^{-i\beta_m\hat{\mathcal{H}}_x}\mathrm{e}^{-i\gamma_m\hat{\mathcal{H}}_{zz}} \;=\; \mathcal{R}_{\hat{\boldsymbol{z}}}(4\beta_m)\mathcal{R}_{\boldsymbol{b}_k}(4\gamma_m)\hat{\boldsymbol{\tau}}_k \tag{A.41}$$

where $R_{\hat{\boldsymbol{\omega}}}(\theta)$ is the 3×3 matrix associated with a rotation of an angle θ around the unit vector $\hat{\boldsymbol{\omega}}$. Composing all the rotations appearing in the definition of $\widehat{U}^{\dagger}_{\rm digit}(\boldsymbol{\gamma},\boldsymbol{\beta})$, see Eq. (4.3), one gets that the final pseudo-spin magnetization $\boldsymbol{\tau}_k(\boldsymbol{\gamma},\boldsymbol{\beta})$ is

$$\begin{aligned}\boldsymbol{\tau}_k(\boldsymbol{\gamma},\boldsymbol{\beta}) \;&=\; \langle\psi_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})|\hat{\boldsymbol{\tau}}_k|\psi_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle = \langle\psi_0|\widehat{U}^{\dagger}_{\rm digit}(\boldsymbol{\gamma},\boldsymbol{\beta})\hat{\boldsymbol{\tau}}_k\widehat{U}_{\rm digit}(\boldsymbol{\gamma},\boldsymbol{\beta})|\psi_0\rangle \\ &=\; \left(\prod_{m=1}^{\leftarrow{\rm P}}\mathcal{R}_{\hat{\boldsymbol{z}}}(4\beta_m)\mathcal{R}_{\boldsymbol{b}_k}(4\gamma_m)\right)\hat{\boldsymbol{z}}\;.\end{aligned} \tag{A.42}$$

Eq. (A.42) holds both when considering a reduced spin chain with PBC or ABC. However, since $\mathcal{K}^{(+)}$ and $\mathcal{K}^{(-)}$ are not equal, the pseudo-spin wave-vectors that contribute to the residual energy $\epsilon_{\mathrm{P}}^{\mathrm{res}}$ depend on the boundary condition.

Indeed, using the first line of Eq. (A.38) and Eqs. (A.33), (A.36),(A.42), the residual energy can be written as

$$\begin{aligned}\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\boldsymbol{\beta}) &= \langle\psi_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta})|\frac{\hat{\mathcal{H}}_z^{(\pm)}+2N_{\mathrm{R}}}{2N_{\mathrm{R}}}|\psi_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle \\ &= \frac{1}{2}-\frac{2}{2N_{\mathrm{R}}}\sum_{k}^{\mathcal{K}^{(\pm)}}\boldsymbol{\tau}_k(\boldsymbol{\gamma},\boldsymbol{\beta})\cdot\hat{\boldsymbol{b}}_k \\ &= \frac{N_{\mathrm{R}}-2|\mathcal{K}^{(\pm)}|}{2N_{\mathrm{R}}}+\frac{1}{2N_{\mathrm{R}}}\sum_{k}^{\mathcal{K}^{(\pm)}}\left\|\boldsymbol{\tau}_k(\boldsymbol{\gamma},\boldsymbol{\beta})-\hat{\boldsymbol{b}}_k\right\|^2 , \end{aligned} \tag{A.43}$$

where in the last step we used that $\hat{\boldsymbol{b}}_k$ and $\boldsymbol{\tau}_k$ are unit vectors, and denoted by $|\mathcal{K}^{(\pm)}|$ the number of k-vectors in $\mathcal{K}^{(\pm)}$.

We can now return to consider a full spin chain with PBC. We recall that, in Chapter 4, we showed that for $2\mathrm{P}+2\leq N$, changing the boundary conditions of the reduced chain does not affect the value of the residual energy. We also showed that using ABC for the reduced chain is indeed convenient in establishing a non-trivial bound for the residual energy. Choosing ABC here, recalling that $2|\mathcal{K}^{(-)}|=N_{\mathrm{R}}-2$, we conclude that for $2\mathrm{P}<N$ we get:

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\boldsymbol{\beta})\overset{2\mathrm{P}<N}{=}\frac{1}{2P+2}+\frac{1}{2P+2}\sum_{k}^{\mathcal{K}^{(-)}}\frac{\left\|\boldsymbol{\tau}_k(\boldsymbol{\gamma},\boldsymbol{\beta})-\hat{\boldsymbol{b}}_k\right\|^2}{2} . \tag{A.44}$$

For $2\mathrm{P}\geq N$ we must use PBC, hence $2|\mathcal{K}^{(+)}|=N_{\mathrm{R}}$, and we get:

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma},\boldsymbol{\beta})\overset{2\mathrm{P}\geq N}{=}\frac{1}{N}\sum_{k}^{\mathcal{K}^{(+)}}\frac{\left\|\boldsymbol{\tau}_k(\boldsymbol{\gamma},\boldsymbol{\beta})-\hat{\boldsymbol{b}}_k\right\|^2}{2} . \tag{A.45}$$

These are the same expressions presented in Eq. (4.20) and Eq. (4.21), as one can immediately show that the expression for $\epsilon_k(\boldsymbol{\gamma},\boldsymbol{\beta})$ in Eq. (4.22) is indeed:

$$\epsilon_k(\boldsymbol{\gamma},\boldsymbol{\beta})=\frac{\left\|\boldsymbol{\tau}_k(\boldsymbol{\gamma},\boldsymbol{\beta})-\hat{\boldsymbol{b}}_k\right\|^2}{2} . \tag{A.46}$$

Appendix B

Asymptotic density of defects for linear QA

In this Appendix we consider a translational invariant quantum Ising chain. We derive an analytical expression for the asymptotic behavior of the density of defects for both a linear *step* -QA and linear *digitized* -QA protocols considered in Chapter 3.

We use Eqs. (3.2) (3.16) (3.17) to write the interpolating QA Hamiltonian as

$$\widehat{H}(s) = s\widehat{H}_z + (1-s)\widehat{H}_x = -\sum_{k>0}^{\text{ABC}} \hbar\boldsymbol{\omega}(s,\cos k)\cdot\frac{\hat{\boldsymbol{\tau}}_k}{2}$$

where we defined $\boldsymbol{\omega}(s,u=\cos k)$ to be the three dimensional vector

$$\begin{aligned}\hbar\boldsymbol{\omega}(s,u) &= 4Js\hat{\boldsymbol{b}}_k + 4\Gamma(1-s)\hat{\boldsymbol{z}} \\ &= 4(Js\sqrt{1-u^2},0,\Gamma(1-s)-Jsu)^T \end{aligned} \tag{B.1}$$

with $\hat{\boldsymbol{z}} = (0,0,1)^T$ and $\hat{\boldsymbol{b}}_k = (\sin k, 0, -\cos k)^T$ as in Chapter 3. The unitary gates in Eqs. (3.5) and (3.10) rotate each pseudo-spin $\hat{\boldsymbol{\tau}}$ around a fixed axis, which depends on the momentum k and possibly on s. In particular, we will use the following Pauli matrix identity:

$$\mathrm{e}^{-i\frac{\theta}{2}\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{\tau}}}\hat{\boldsymbol{\tau}}\mathrm{e}^{+i\frac{\theta}{2}\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{\tau}}} = \mathcal{R}_{\hat{\boldsymbol{n}}}(\theta)\hat{\boldsymbol{\tau}}\ , \tag{B.2}$$

where $\mathcal{R}_{\hat{\boldsymbol{n}}}(\theta)$ is a 3×3 rotation matrix around the axis $\hat{\boldsymbol{n}}$ by an angle θ, acting on the pseudo-spin Cartesian components. It will be useful, in the following, to have an explicit expression for such a rotation matrix for generic $\hat{\boldsymbol{n}}$ and θ. For that

purpose, we recast it in the form $\mathcal{R}_{\hat{\boldsymbol{n}}}(\theta) = \mathbb{R}[\boldsymbol{\Omega} = \theta\hat{\boldsymbol{n}}]$, where $\Omega = \sqrt{\boldsymbol{\Omega}\cdot\boldsymbol{\Omega}} = \theta$ and:

$$\begin{aligned}(\mathbb{R}[\boldsymbol{\Omega}])_{ij} &= \frac{\Omega_i\Omega_j}{\Omega^2} + \left(\delta_{ij} - \frac{\Omega_i\Omega_j}{\Omega^2}\right)\cos\Omega \\ &\quad + \left(\sum_{k=1}^{3}\epsilon_{ijk}\frac{\Omega_k}{\Omega}\right)\sin\Omega\ . \end{aligned} \tag{B.3}$$

To perform a *step*-QA dynamics, the typical ingredient needed would be:

$$\mathrm{e}^{\frac{i\Delta t}{\hbar}\widehat{H}(s_m)}\hat{\boldsymbol{\tau}}_k\mathrm{e}^{-\frac{i\Delta t}{\hbar}\widehat{H}(s_m)} = \mathbb{R}[\boldsymbol{\omega}(s_m,\cos k)\,\Delta t]\hat{\boldsymbol{\tau}}_k\ . \tag{B.4}$$

In a *digitized*-QA, the ingredient needed is:

$$\widehat{U}_m^\dagger\hat{\boldsymbol{\tau}}_k\widehat{U}_m = \mathbb{R}[4\beta_m\Gamma\hat{\boldsymbol{z}}]\,\mathbb{R}[4\gamma_m J\hat{\boldsymbol{b}}_k]\hat{\boldsymbol{\tau}}_k\ , \tag{B.5}$$

where $\widehat{U}_m = \mathrm{e}^{-i\beta_m\widehat{H}_x}\mathrm{e}^{-i\gamma_m\widehat{H}_z}$ and one should observe the order of the rotation matrices applied. Recall now that the density of defects, see Eq. (3.20), is expressed in the thermodynamic limit as:

$$\rho_{\mathrm{def}}(\tau) = \frac{1}{2} - \int_0^\pi \frac{\mathrm{d}k}{2\pi}\,\hat{\boldsymbol{b}}_k\cdot\langle\psi_0|\widehat{U}^\dagger(\tau)\hat{\boldsymbol{\tau}}_k\widehat{U}(\tau)|\psi_0\rangle\ . \tag{B.6}$$

For the step-QA case the quantum average needed is:

$$\begin{aligned}\langle\hat{\boldsymbol{\tau}}_k\rangle_\tau^{\mathrm{step}} &= \langle\psi_0|\widehat{U}_{\mathrm{step}}^\dagger(\tau)\hat{\boldsymbol{\tau}}_k\widehat{U}_{\mathrm{step}}(\tau)|\psi_0\rangle \\ &= \left[\prod_{m=1}^{\overleftarrow{\mathrm{P}}}\mathbb{R}[\boldsymbol{\omega}(s_m,\cos k)\tfrac{\tau}{\mathrm{P}}]\right]\langle\psi_0|\hat{\boldsymbol{\tau}}_k|\psi_0\rangle\ , \end{aligned} \tag{B.7}$$

where $\prod^{\overleftarrow{\mathrm{P}}}$ denotes a time-ordered product, and $|\psi_0\rangle$ is the initial state of the system. Since $|\psi_0\rangle$ is chosen to be the ground state of the initial Hamiltonian $\widehat{H}(0) = \widehat{H}_x$ we have:

$$\widehat{H}(0) = -2\Gamma\sum_{k>0}^{\mathrm{ABC}}\hat{\boldsymbol{z}}\cdot\hat{\boldsymbol{\tau}}_k \implies \langle\psi_0|\hat{\boldsymbol{\tau}}_k|\psi_0\rangle = \hat{\boldsymbol{z}}\ . \tag{B.8}$$

This leads to the equations:

$$\langle \hat{\boldsymbol{\tau}}_k \rangle_\tau^{\text{step}} = \left[\prod_{m=1}^{\leftarrow \mathrm{P}} \mathbb{R}[\boldsymbol{\omega}(s_m, \cos k) \tfrac{\tau}{\mathrm{P}}] \right] \hat{\boldsymbol{z}} \tag{B.9}$$

$$\langle \hat{\boldsymbol{\tau}}_k \rangle_\tau^{\text{digit}} = \left[\prod_{m=1}^{\leftarrow \mathrm{P}} \mathbb{R}[4\beta_m \Gamma \hat{\boldsymbol{z}}] \, \mathbb{R}[4\gamma_m J \hat{\boldsymbol{b}}_k] \right] \hat{\boldsymbol{z}} \, . \tag{B.10}$$

It is convenient to extract the asymptotic defect density $\rho_{\text{def}}(\tau \to \infty)$ from the infinite-time average:

$$\overline{\rho}_{\text{def}} = \lim_{\mathrm{T}\to\infty} \frac{1}{\mathrm{T}} \int_0^{\mathrm{T}} \mathrm{d}\tau \, \rho_{\text{def}}(\tau) \, . \tag{B.11}$$

In the thermodynamic limit, exchanging the k-integral with the time-integral we get:

$$\overline{\rho}_{\text{def}} = \frac{1}{2} - \int_0^{\pi} \frac{\mathrm{d}k}{2\pi} \, \hat{\boldsymbol{b}}_k \cdot \overline{\langle \hat{\boldsymbol{\tau}}_k \rangle_\tau} \, . \tag{B.12}$$

From now on, we simplify our notation by adopting units such that $\hbar = 1$ and $J = 1$. We will also take $\Gamma = 1$, so that the Ising critical point is located at $s_c = \frac{1}{2}$. With this choice of units we have that $\boldsymbol{\omega}(1, \cos k) = 4\hat{\boldsymbol{b}}_k$ and $\boldsymbol{\omega}(0, \cos k) = 4\hat{\boldsymbol{z}}$.

B.1 Linear step-QA analysis

For the *step*-QA evolution we get

$$\begin{aligned} \overline{\langle \hat{\boldsymbol{\tau}}_k \rangle}_\tau^{\text{step}} &= \lim_{\mathrm{T}\to\infty} \frac{1}{\mathrm{T}} \int_0^{\mathrm{T}} \mathrm{d}\tau \, \langle \hat{\boldsymbol{\tau}}_k \rangle_\tau^{\text{step}} \\ &= \lim_{\mathrm{T}\to\infty} \frac{1}{\mathrm{T}} \int_0^{\mathrm{T}} \mathrm{d}\tau \left[\prod_{m=1}^{\leftarrow \mathrm{P}} \mathbb{R}[\boldsymbol{\omega}(s_m, \cos k) \tfrac{\tau}{\mathrm{P}}] \right] \hat{\boldsymbol{z}} \, , \end{aligned}$$

which leads to

$$\begin{aligned} \overline{\rho}_{\text{def,P}}^{\text{step}} &= \frac{1}{2} - \frac{1}{8} \int_{-\pi}^{\pi} \frac{\mathrm{d}k}{2\pi} \, \boldsymbol{\omega}(1, \cos k) \cdot \overline{\langle \hat{\boldsymbol{\tau}}_k \rangle}_\tau^{\text{step}} = \\ &= \frac{1}{2} - \frac{1}{32} \oint_C \frac{\mathrm{d}z}{2\pi i z} \, \boldsymbol{\omega}_{1,z}^T \lim_{\mathrm{T}\to\infty} \frac{1}{\mathrm{T}} \int_0^{\mathrm{T}} \mathrm{d}\tau \left[\prod_{m=1}^{\leftarrow \mathrm{P}} \mathbb{R}[\boldsymbol{\omega}_{s_m,z} \tau] \right] \boldsymbol{\omega}_{0,z} \, , \end{aligned} \tag{B.13}$$

where we changed variable to $z = e^{ik}$ (so that $\cos k = \frac{z+z^{-1}}{2}$, $\sin k = \frac{z-z^{-1}}{2i}$) and defined:

$$\boldsymbol{\omega}_{s,z} \stackrel{\text{def}}{=} \boldsymbol{\omega}(s, \tfrac{z+z^{-1}}{2}) \, . \tag{B.14}$$

C denotes the unit circle in the complex plane.

As we will show in the following, for any positive integer P, the integrand is a rational function of z and the integral can be evaluated using the residue theorem. To show this we start by observing that the frequencies involved in the time average are

$$\omega^2(s_m, u) = 16[1 - 2s_m(1 - s_m)(u + 1)] \,. \tag{B.15}$$

These are commensurate (indeed identical) only for pairs that are symmetric with respect to $m = \frac{\mathrm{P}}{2}$ (assuming P to be even) because $s_m = 1 - s_{\mathrm{P}-m}$, hence $\omega(s_m, u) = \omega(s_{\mathrm{P}-m}, u)$. Therefore in computing the time average of the product, we may neglect all correlations except those between the symmetric pairs $\mathbb{R}[\boldsymbol{\omega}_{s_m,z}\tau]$ and $\mathbb{R}[\boldsymbol{\omega}_{s_{\mathrm{P}-m},z}\tau]$. The central matrix $\mathbb{R}[\boldsymbol{\omega}_{s_{\frac{\mathrm{P}}{2}},z}\tau]$ and the leftmost matrix $\mathbb{R}[\boldsymbol{\omega}_{s_{\mathrm{P}},z}\tau]$ are unpaired and should be averaged separately. Here is a scheme of the averages we need to perform:

$$\overline{\mathbb{R}[\boldsymbol{\omega}_{s_{\mathrm{P}},z}\tau]}\,\mathbb{R}[\boldsymbol{\omega}_{s_{\mathrm{P}-1},z}\tau]\cdots\mathbb{R}[\boldsymbol{\omega}_{s_{\frac{\mathrm{P}}{2}+1},z}\tau]\,\overline{\mathbb{R}[\boldsymbol{\omega}_{s_{\frac{\mathrm{P}}{2}},z}\tau]}\,\mathbb{R}[\boldsymbol{\omega}_{s_{\frac{\mathrm{P}}{2}-1},z}\tau]\cdots\mathbb{R}[\boldsymbol{\omega}_{s_1,z}\tau] \,. \tag{B.16}$$

To exploit this structure of nested averages, it is convenient to recursively contract (performing the time-integration) two rotation matrices sandwiching the central matrix at $m = \frac{\mathrm{P}}{2}$. By doing this, at each step we can independently average over τ. To carry out the contractions, it is useful to define a rescaled vector:

$$\begin{aligned} \boldsymbol{\Omega}_{s,z} &= z\boldsymbol{\omega}_{s,z} \\ &= 2\left(-is(z^2-1), 0, 2(1-s)z - s(z^2+1)\right)^T \end{aligned} \tag{B.17}$$

and two 3×3 matrices:

$$\mathbb{P}_{s,z} = \boldsymbol{\Omega}_{s,z}\boldsymbol{\Omega}_{s,z}^T \quad \text{and} \quad \mathbb{Q}_{s,z} = -\boldsymbol{\Omega}_{s,z}\wedge \,, \tag{B.18}$$

where $\mathbb{P}_{s,z}$ and $\mathbb{Q}_{s,z}$ should be regarded as acting on vectors as follows: $\mathbb{P}_{s,z}\boldsymbol{a} = \boldsymbol{\Omega}_{s,z}(\boldsymbol{\Omega}_{s,z}\cdot\boldsymbol{a})$ and $\mathbb{Q}_{s,z}\boldsymbol{a} = \boldsymbol{a}\wedge\boldsymbol{\Omega}_{s,z}$. Notice that $\boldsymbol{\Omega}_{s,z}$, $\mathbb{P}_{s,z}$ and $\mathbb{Q}_{s,z}$ are polynomials in z, and they are crucial ingredients appearing in the rotation matrix in Eq. (B.3):

$$\mathbb{R}[\boldsymbol{\omega}_{s,z}\tau] = \frac{1}{\Omega_{s,z}^2}\mathbb{P}_{s,z} + \left(\mathbb{1} - \frac{1}{\Omega_{s,z}^2}\mathbb{P}_{s,z}\right)\cos(\omega_{s,z}\tau) + \frac{1}{\Omega_{s,z}}\mathbb{Q}_{s,z}\sin(\omega_{s,z}\tau) \,. \tag{B.19}$$

The time-average of the central matrix $\mathbb{R}[\boldsymbol{\omega}_{\frac{1}{2},z}\tau]$ is therefore given by:

$$\overline{\mathbb{R}[\boldsymbol{\omega}_{\frac{1}{2},z}\tau]} = \lim_{\mathrm{T}\to\infty}\frac{1}{\mathrm{T}}\int_0^{\mathrm{T}}\mathrm{d}\tau\,\mathbb{R}[\boldsymbol{\omega}_{\frac{1}{2},z}\tau] = \frac{1}{\Omega^2_{\frac{1}{2},z}}\mathbb{P}_{\frac{1}{2},z} \tag{B.20}$$

since only the first term in Eq. (B.19) contributes. A similar expression holds for the average of the leftmost matrix. To recursively contract the terms as indicated in Eq. (B.16), we now define a super-operator $\mathcal{L}_{s,z}$ performing the time-integration of two matrices sandwiching a central term $\mathbb{A}$ (a 3×3 matrix originating from the previous step) as follows:

$$\mathcal{L}_{s,z}\mathbb{A} \stackrel{\text{def}}{=} \Omega^4_{s,z}\mathbb{R}[\overline{\boldsymbol{\omega}_{1-s,z}\tau}]\,\mathbb{A}\,\mathbb{R}[\boldsymbol{\omega}_{s,z}\tau] \stackrel{\text{def}}{=} \Omega^4_{s,z}\lim_{\mathrm{T}\to\infty}\frac{1}{\mathrm{T}}\int_0^{\mathrm{T}}\mathrm{d}\tau\,\mathbb{R}[\boldsymbol{\omega}_{1-s,z}\tau]\,\mathbb{A}\,\mathbb{R}[\boldsymbol{\omega}_{s,z}\tau] \tag{B.21}$$

The time-integral can be easily calculated by exploiting the explicit form of the rotation matrices in Eq. (B.19):

$$\mathcal{L}_{s,z}\mathbb{A} = \frac{1}{2}\Omega^4_{s,z}\mathbb{A} + \frac{1}{2}\Omega^2_{s,z}(\mathbb{Q}_{1-s,z}\mathbb{A}\mathbb{Q}_{s,z} - \mathbb{A}\mathbb{P}_{s,z} - \mathbb{P}_{1-s,z}\mathbb{A}) + \frac{3}{2}\mathbb{P}_{1-s,z}\mathbb{A}\mathbb{P}_{s,z}\ , \tag{B.22}$$

where we used simple trigonometric integrals, such as $\overline{\cos^2(\omega_{s,z}\tau)} = \overline{\sin^2(\omega_{s,z}\tau)} = \frac{1}{2}$, and the fact that $\Omega_{1-s,z} = \Omega_{s,z}$. With this device, we can write an explicit expression for the time-averaged defect density in terms of a time-ordered product of $\frac{\mathrm{P}}{2}-1$ super-operators

$$\overline{\rho}^{\text{step}}_{\text{def},\mathrm{P}} = \frac{1}{2} - \frac{1}{32}\oint_C\frac{\mathrm{d}z}{2\pi i z}\,\frac{\boldsymbol{\Omega}^T_{1,z}\left[\left(\prod_{m=\frac{\mathrm{P}}{2}+1}^{\leftarrow \mathrm{P}-1}\mathcal{L}_{1-s_m,z}\right)\mathbb{P}_{\frac{1}{2},z}\right]\boldsymbol{\Omega}_{0,z}}{z^2\left(\prod_{m=\frac{\mathrm{P}}{2}+1}^{\mathrm{P}-1}\Omega^4_{s_m,z}\right)\Omega^2_{\frac{1}{2},z}}\ , \tag{B.23}$$

where the (unpaired) leftmost rotation has been treated separately and did not contribute.

The polynomial appearing in the denominator can be factorized using

$$\Omega^4_{s,z} = 2^8 s^2(1-s)^2 z^2\left(z-\frac{1-s}{s}\right)^2\left(z-\frac{s}{1-s}\right)^2\ . \tag{B.24}$$

Moreover, we have:

$$\frac{1}{\Omega^2_{\frac{1}{2},z}}\mathbb{P}_{\frac{1}{2},z} = -\frac{1}{4z}\widetilde{\mathbb{P}}_z\ , \tag{B.25}$$

where $\widetilde{\mathbb{P}}_z = \widetilde{\mathbf{\Omega}}_z \widetilde{\mathbf{\Omega}}_z^T$ and $\widetilde{\mathbf{\Omega}}_z = (-i(z+1), 0, -(z-1))^T$. We therefore rewrite our final expression for $\overline{\rho}^{\rm step}_{\rm def,P}$ as the complex integral of a rational function of z:

$$\overline{\rho}^{\rm step}_{\rm def,P} = \frac{1}{2} + \oint_C \frac{{\rm d}z}{2\pi i} \frac{f_{\rm P}(z)}{g_{\rm P}(z)} \,, \tag{B.26}$$

where $f_{\rm P}(z)$ and $g_{\rm P}(z)$ are polynomials of z:

$$f_{\rm P}(z) = \mathbf{\Omega}_{1,z}^T \left[\left(\prod_{m=\frac{\rm P}{2}+1}^{\leftarrow {\rm P}-1} \mathcal{L}_{1-s_m,z} \right) \widetilde{\mathbb{P}}_z \right] \mathbf{\Omega}_{0,z} \tag{B.27}$$

$$\begin{aligned} g_{\rm P}(z) &= 2^{4{\rm P}-1} z^{{\rm P}+2} \prod_{m=\frac{\rm P}{2}+1}^{{\rm P}-1} \frac{m^2}{{\rm P}^2} \left(1-\frac{m}{\rm P}\right)^2 \left(z - \frac{{\rm P}-m}{m}\right)^2 \left(z - \frac{m}{{\rm P}-m}\right)^2 \\ &= 2^{4{\rm P}-1} z^{\rm P} \prod_{m=\frac{\rm P}{2}+1}^{\rm P} \frac{m^2}{{\rm P}^2} \left(1-\frac{m}{\rm P}\right)^2 \left(z - \frac{{\rm P}-m}{m}\right)^2 \left(z - \frac{m}{{\rm P}-m}\right)^2 \end{aligned} \tag{B.28}$$

Here, in the last expression for $g_{\rm P}(z)$ we have re-expressed a factor z^2 has an extra term in the product, with $m = {\rm P}$. This expression for $g_{\rm P}(z)$ shows a number of poles of the rational function which will be used in the calculation of residues.

As an example, for the case ${\rm P} = 2$ we have $f_{\rm P=2}(z) = 2^4 z^2 (z-1)^2$ and $g_{\rm P=2}(z) = 2^7 z^4$, hence:

$$\begin{aligned} \overline{\rho}^{\rm step}_{\rm def,P=2} &= \frac{1}{2} + \oint_C \frac{{\rm d}z}{2\pi i} \frac{f_2(z)}{g_2(z)} \\ &= \frac{1}{2} + \oint_C \frac{{\rm d}z}{2\pi i} \frac{(z-1)^2}{8z^2} = \frac{1}{4} \,. \end{aligned} \tag{B.29}$$

For ${\rm P} = 4$ we get:

$$\begin{aligned} g_{\rm P=4}(z) &= 9 \cdot 2^7 z^6 (z-\tfrac{1}{3})^2 (z-3)^2 \\ f_{\rm P=4}(z) &= 2^3 z^4 (z-1)^2 (9z^4 - 92z^3 + 310z^2 - 92z + 9) \,, \end{aligned}$$

hence:

$$\overline{\rho}^{\rm step}_{\rm def,P=4} = \frac{1}{2} + \frac{1}{2\pi i} \oint_C {\rm d}z \frac{f_4(z)}{g_4(z)} = \frac{13}{72} \,. \tag{B.30}$$

More in general, we can use the residue theorem to calculate:

$$\overline{\rho}^{\rm step}_{\rm def,P} = \frac{1}{2} + \sum_{m=0}^{\frac{\rm P}{2}-1} \frac{(z-z_m)[f_{\rm P}(z)-f_{\rm P}(z_m)]}{g_{\rm P}(z)}\Big|_{z=z_m} , \tag{B.31}$$

where $z_m = \frac{m}{\rm P-m}$ for $m = 0, 1 \ldots (\frac{\rm P}{2}-1)$ are the $\frac{\rm P}{2}$ (double) roots of the polynomial $g_{\rm P}(z)$ lying inside the unitary circle C.

For higher values of P, the polynomials can be calculated with Mathematica, and integrals then evaluated with the residue theorem.

B.2 Linear digitized-QA analysis

For the *digitized*-QA evolution we get:

$$\begin{aligned} \overline{\langle \hat{\boldsymbol{\tau}}_k \rangle_\tau}^{\rm digit} &= \lim_{\rm T\to\infty} \frac{1}{\rm T} \int_0^{\rm T} {\rm d}\tau \, \langle \hat{\boldsymbol{\tau}}_k \rangle_\tau^{\rm digit} \\ &= \lim_{\rm T\to\infty} \frac{1}{\rm T} \int_0^{\rm T} {\rm d}\tau \left[\prod_{m=1}^{\leftarrow {\rm P}} \mathbb{R}[4\beta_m \hat{\boldsymbol{z}}] \, \mathbb{R}[4\gamma_m \hat{\boldsymbol{b}}_k] \right] \hat{\boldsymbol{z}} , \end{aligned}$$

which leads, upon the usual change of variables:

$$\begin{aligned} \overline{\rho}^{\rm digit}_{\rm def,P} &= \frac{1}{2} - \frac{1}{8} \int_{-\pi}^{\pi} \frac{{\rm d}k}{2\pi} \, \boldsymbol{\omega}(1, \cos k) \cdot \overline{\langle \hat{\boldsymbol{\tau}}_k \rangle_\tau}^{\rm digit} = \\ &= \frac{1}{2} - \frac{1}{32} \oint_C \frac{{\rm d}z}{2\pi i z} \, \boldsymbol{\omega}_{1,z}^T \lim_{\rm T\to\infty} \frac{1}{\rm T} \int_0^{\rm T} {\rm d}\tau \Bigg[\\ &\qquad \prod_{m=1}^{\leftarrow {\rm P}} \mathbb{R}[(1 - \tfrac{s_m + s_{m+1}}{2}) \boldsymbol{\omega}_{0,z} \tau] \, \mathbb{R}[s_m \boldsymbol{\omega}_{1,z} \tau] \Bigg] \boldsymbol{\omega}_{0,z} , \end{aligned} \tag{B.32}$$

where we made use of the symmetric Trotter splitting and rescaled variables $\tau/{\rm P} \to \tau$ in the time-integral. Recall we should assume that $s_m = m/{\rm P}$ with $s_{{\rm P}+1} \equiv 1$, in order to make $\beta_{\rm P} = 0$. The frequencies appearing in the various rotation matrices are now all *different*, as one can easily verify. The time average of the product then becomes a product of time-averages:

$$\lim_{\rm T\to\infty} \frac{1}{\rm T} \int_0^{\rm T} {\rm d}\tau \left[\prod_{m=1}^{\leftarrow {\rm P}} \mathbb{R}[(1 - \tfrac{s_m + s_{m+1}}{2}) \boldsymbol{\omega}_{0,z} \tau] \, \mathbb{R}[s_m \boldsymbol{\omega}_{1,z} \tau] \right] = \left[\frac{1}{\Omega_{0,z}^2 \Omega_{1,z}^2} \mathbb{P}_{0,z} \mathbb{P}_{1,z} \right]^{\rm P} . \tag{B.33}$$

where, we recall that:

$$\begin{aligned} \mathbf{\Omega}_{0,z} &= z\boldsymbol{\omega}_{0,z} = 4z\,(0,0,1)^T \\ \mathbf{\Omega}_{1,z} &= z\boldsymbol{\omega}_{1,z} = 2\left(-i(z^2-1),0,-(z^2+1)\right)^T \ , \end{aligned} \tag{B.34}$$

which imply:

$$\begin{aligned} \Omega_{0,z}^2 &= \mathbf{\Omega}_{0,z}\cdot\mathbf{\Omega}_{0,z} = 16z^2 \\ \Omega_{1,z}^2 &= \mathbf{\Omega}_{1,z}\cdot\mathbf{\Omega}_{1,z} = 16z^2 \ . \end{aligned} \tag{B.35}$$

Since it will be useful in a moment, we also calculate:

$$\mathbf{\Omega}_{1,z}\cdot\mathbf{\Omega}_{0,z} = -8z(z^2+1) \ . \tag{B.36}$$

The final expression can therefore be cast in the form:

$$\begin{aligned} \overline{\rho}_{\mathrm{def,P}}^{\mathrm{digit}} &= \frac{1}{2} - \frac{1}{32}\oint_C \frac{\mathrm{d}z}{2\pi i z}\,\frac{\mathbf{\Omega}_{1,z}^T\left(\mathbb{P}_{0,z}\mathbb{P}_{1,z}\right)^{\mathrm{P}}\mathbf{\Omega}_{0,z}}{z^2(16z^2)^{2\mathrm{P}}} \\ &= \frac{1}{2} - \frac{1}{2^{8\mathrm{P}+5}}\oint_C \frac{\mathrm{d}z}{2\pi i z}\,\frac{\left(\mathbf{\Omega}_{1,z}\cdot\mathbf{\Omega}_{0,z}\right)^{\mathrm{P}+1}}{z^{4\mathrm{P}+2}} \\ &= \frac{1}{2} + \frac{(-1)^{\mathrm{P}}}{2^{5\mathrm{P}+2}}\oint_C \frac{\mathrm{d}z}{2\pi i}\,\frac{(z^2+1)^{\mathrm{P}+1}}{z^{3\mathrm{P}+2}} \ . \end{aligned} \tag{B.37}$$

We now observe that for even P only even powers of z appear inside the integral, and the residue vanishes. For $\mathrm{P} > 1$ and odd, the numerator is a polynomial with maximum degree $z^{2\mathrm{P}+2}$, hence once again the residue vanishes. The only case in which the integral gives a contribution is for $\mathrm{P} = 1$, where we get:

$$\overline{\rho}_{\mathrm{def,P=1}}^{\mathrm{digit}} = \frac{1}{2} - \frac{1}{2^7}\oint_C \frac{\mathrm{d}z}{2\pi i}\,\frac{(z^2+1)^2}{z^5} = \frac{1}{2} - \frac{1}{2^7} = \frac{63}{128} \ ,$$

which is very close to 1/2 and probably indistinguishable from it, given the fact that there are large fluctuations around the average (which we do not study here).

Appendix C

Translational invariance with mixed boundary conditions

One of the key points in proving the variational bounds given in Chapters 4 and 5 is the possibility to "restore" translational invariance in spin systems with mixed BC (where both ABC and PBC allowed). In this Section, we discuss this property and the associated identities used in Chapters 4 and 5.

In Sec. C.1 we give a simple proof in the specific case of the quantum Ising chain considered in Chapter 4. Later, in Sec. C.2, the proof is generalized to the more general setting considered in Chapter 5.

C.1 Translational invariant quantum Ising chain

We consider a reduced quantum Ising chain of $N_{\rm R} = 2{\rm P} + 2$ spins with anti-periodic boundary conditions (ABC). As in Chapter 4, we number the spins using their position relative to the reduced chain. In this section we give a proof of the identity

$$\langle\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})|\hat{\sigma}^z_{j_s}\hat{\sigma}^z_{j_s+1}|\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle \;=\; \frac{1}{N_{\rm R}}\langle\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})|\Big(\sum_{j=1}^{N_{\rm R}-1}\hat{\sigma}^z_j\hat{\sigma}^z_{j+1} - \hat{\sigma}^z_{N_{\rm R}}\hat{\sigma}^z_1\Big)|\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle \tag{C.1}$$

where j_s is any internal lattice site, and $|\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle$ (see Eq. (4.15)) is

$$|\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle = \prod_{m=1}^{\leftarrow{\rm P}} {\rm e}^{-i\hat{\mathcal{H}}_x\beta_m}{\rm e}^{-i\hat{\mathcal{H}}_z^{(-)}\gamma_m}\,|\widetilde{\psi}_0\rangle\;, \tag{C.2}$$

with $|\widetilde{\psi}_0\rangle = |+\rangle^{\otimes N_{\rm R}}$ and the reduced chain Hamiltonians given by:

$$\hat{\mathcal{H}}_z^{(-)} + N_{\rm R} = \sum_{j=1}^{N_{\rm R}-1} \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z - \hat{\sigma}_{N_{\rm R}}^z \hat{\sigma}_1^z \tag{C.3}$$

$$\hat{\mathcal{H}}_x = -\sum_{j=1}^{N_{\rm R}} \hat{\sigma}_j^x \,. \tag{C.4}$$

Notice that the expression appearing on the right-hand side of Eq. (C.1) coincides, apart from the constant $N_{\rm R}$, with $\hat{\mathcal{H}}_z^{(-)}$, the reduced spin chain Hamiltonian with ABC ($J_b = -1$) introduced in Eq. (4.12). In the main text we used the identity in Eq. (C.1) to derive the expression for the residual energy given in Eq. (4.18).

The key to the proof of Eq. (C.1) is showing that there exits a unitary "anti-periodic" translation transformation $\widehat{\mathcal{T}}_{\rm ABC}$ that is a symmetry of the Hamiltonians. [1] Given the usual translation operator $\widehat{\mathcal{T}}_{\rm PBC}$:

$$\widehat{\mathcal{T}}_{\rm PBC}^\dagger \hat{\boldsymbol{\sigma}}_j \widehat{\mathcal{T}}_{\rm PBC} = \hat{\boldsymbol{\sigma}}_{j+1} \qquad \text{for } j \neq N_{\rm R} \tag{C.5}$$

$$\widehat{\mathcal{T}}_{\rm PBC}^\dagger \hat{\boldsymbol{\sigma}}_{N_{\rm R}} \widehat{\mathcal{T}}_{\rm PBC} = \hat{\boldsymbol{\sigma}}_1 \,, \tag{C.6}$$

we define the anti-periodic translation operator $\widehat{\mathcal{T}}_{\rm ABC}$ to be the unitary transformation obtained by composing the standard translation $\widehat{\mathcal{T}}_{\rm PBC}$ with a flip of the first spin: $\widehat{\mathcal{T}}_{\rm ABC} \equiv \widehat{\mathcal{T}}_{\rm PBC} \hat{\sigma}_1^x$. The action on the spin operators induced by $\widehat{\mathcal{T}}_{\rm ABC}$ is

$$\widehat{\mathcal{T}}_{\rm ABC}^\dagger \hat{\boldsymbol{\sigma}}_j \widehat{\mathcal{T}}_{\rm ABC} = \hat{\sigma}_1^x \hat{\boldsymbol{\sigma}}_{j+1} \hat{\sigma}_1^x = \hat{\boldsymbol{\sigma}}_{j+1} \qquad \text{for } j \neq N_{\rm R} \tag{C.7}$$

$$\widehat{\mathcal{T}}_{\rm ABC}^\dagger \hat{\boldsymbol{\sigma}}_{N_{\rm R}} \widehat{\mathcal{T}}_{\rm ABC} = \hat{\sigma}_1^x \hat{\boldsymbol{\sigma}}_1 \hat{\sigma}_1^x = (\hat{\sigma}_1^x, -\hat{\sigma}_1^y, -\hat{\sigma}_1^z)^T \,. \tag{C.8}$$

Using Eq. (C.7) and Eq. (C.8), a straightforward computation shows that

$$\widehat{\mathcal{T}}_{\rm ABC}^\dagger \hat{\mathcal{H}}_z^{(-)} \widehat{\mathcal{T}}_{\rm ABC} = \hat{\mathcal{H}}_z^{(-)} \,, \tag{C.9}$$

while the invariance of $\hat{\mathcal{H}}_x$ and of the inital state $|\widetilde{\psi}_0\rangle$ is trivial. This in turns implies the identity

$$\widehat{\mathcal{T}}_{\rm ABC} |\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = |\widetilde{\psi}_{\rm P}(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle \,. \tag{C.10}$$

[1] Later on, in Sec. C.2, and in the whole Chapter 5, the notation used is slightly different: $\widehat{\mathcal{T}}_{\rm ABC} \to \widehat{T}_{\hat{\boldsymbol{x}}}^{(-)}$ and $\widehat{\mathcal{T}}_{\rm PBC} \to \widehat{T}_{\hat{\boldsymbol{x}}}^{(+)}$.

Additionally, Eq. (C.7) and Eq. (C.8) also imply that $\hat{\mathcal{H}}_z^{(-)} + N_{\mathrm{R}}$ decomposes into a sum of terms obtained by applying powers of $\widehat{\mathcal{T}}_{\mathrm{ABC}}$ to $\hat{\sigma}_{j_s}^z \hat{\sigma}_{j_s+1}^z$:

$$\hat{\mathcal{H}}_z^{(-)} + N_{\mathrm{R}} \;=\; \sum_{n=0}^{N_{\mathrm{R}}-1} \widehat{\mathcal{T}}_{\mathrm{ABC}}^{\dagger n}\, \hat{\sigma}_{j_s}^z \hat{\sigma}_{j_s+1}^z\, \widehat{\mathcal{T}}_{\mathrm{ABC}}^{n} \,. \tag{C.11}$$

The desired equality is a direct consequence of Eq. (C.10) and Eq. (C.11). Indeed, one has that

$$\begin{aligned}
\langle \widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta}) | \hat{\mathcal{H}}_z^{(-)} + N_{\mathrm{R}} | \widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta}) \rangle \;&=\; \langle \widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta}) | \sum_{n=0}^{N_{\mathrm{R}}-1} \widehat{\mathcal{T}}_{\mathrm{ABC}}^{\dagger n}\, \hat{\sigma}_{j_s}^z \hat{\sigma}_{j_s+1}^z\, \widehat{\mathcal{T}}_{\mathrm{ABC}}^{n} | \widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta}) \rangle \\
&=\; \sum_{n=0}^{N_{\mathrm{R}}-1} \langle \widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta}) | \hat{\sigma}_{j_s}^z \hat{\sigma}_{j_s+1}^z | \widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta}) \rangle \\
&=\; N_{\mathrm{R}} \langle \widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta}) | \hat{\sigma}_{j_s}^z \hat{\sigma}_{j_s+1}^z | \widetilde{\psi}_{\mathrm{P}}(\boldsymbol{\gamma},\boldsymbol{\beta}) \rangle \,,
\end{aligned} \tag{C.12}$$

which proves Eq. (C.1).

C.2 Translational invariant spin systems

We show here that if $\widehat{T}_{\mathbf{r}}^{(s_x s_y)}$ denote the modified translation operators defined in Chapter 5 (see, for instance, Eq. (5.19)), and the reduced Hamiltonians are defined as:

$$\hat{\mathcal{H}}_{\mu}^{(s_x s_y)} = \sum_{\mathbf{r}\in\mathcal{L}_{\mathrm{R}}} \widehat{T}_{\mathbf{r}}^{(s_x s_y)\dagger}\, \widehat{O}_{\mu}(\mathbf{0})\, \widehat{T}_{\mathbf{r}}^{(s_x s_y)} \,. \tag{C.13}$$

then, automatically:

$$\widehat{T}_{\mathbf{r}}^{(s_x s_y)\dagger}\, \hat{\mathcal{H}}_{\mu}^{(s_x s_y)}\, \widehat{T}_{\mathbf{r}}^{(s_x s_y)} = \hat{\mathcal{H}}_{\mu}^{(s_x s_y)} \,, \tag{C.14}$$

provided the corresponding $\widehat{O}_{\mu}$ is spin-flip symmetric, i.e.:

$$\widehat{\mathcal{P}}^{\dagger} \widehat{O}_{\mu} \widehat{\mathcal{P}} = \widehat{O}_{\mu} \,. \tag{C.15}$$

To prove Eq. (C.14), it is enough to show that $\hat{\mathcal{H}}_{\mu}^{(s_x s_y)}$ commutes with the elementary translations along the two axis

$$\begin{cases}
\widehat{T}_{\hat{\boldsymbol{x}}}^{(s_x)\dagger}\, \hat{\mathcal{H}}_{\mu}^{(s_x s_y)}\, \widehat{T}_{\hat{\boldsymbol{x}}}^{(s_x)} = \hat{\mathcal{H}}_{\mu}^{(s_x s_y)} \\
\\
\widehat{T}_{\hat{\boldsymbol{y}}}^{(s_y)\dagger}\, \hat{\mathcal{H}}_{\mu}^{(s_x s_y)}\, \widehat{T}_{\hat{\boldsymbol{y}}}^{(s_y)} = \hat{\mathcal{H}}_{\mu}^{(s_x s_y)}
\end{cases} , \tag{C.16}$$

and use the commutativity of the various translation operators. We write the explicit expression for $\hat{\boldsymbol{x}}$:

$$\begin{aligned}
\widehat{T}^{(s_x)\dagger}_{\hat{\boldsymbol{x}}}\,\hat{\mathcal{H}}^{(s_x s_y)}_{\mu}\,\widehat{T}^{(s_x)}_{\hat{\boldsymbol{x}}} \;-\; \hat{\mathcal{H}}^{(s_x s_y)}_{\mu} \;&= \\
&= \sum_{n=1}^{L_{\mathrm{R}}} \sum_{m=0}^{L_{\mathrm{R}}-1} \widehat{T}^{(s_y)\dagger}_{m\hat{\boldsymbol{y}}}\,\widehat{T}^{(s_x)\dagger}_{n\hat{\boldsymbol{x}}}\,\widehat{O}_{\mu}(\mathbf{0})\,\widehat{T}^{(s_x)}_{n\hat{\boldsymbol{x}}}\,\widehat{T}^{(s_y)}_{m\hat{\boldsymbol{y}}} \\
&\qquad - \sum_{n=0}^{L_{\mathrm{R}}-1} \sum_{m=0}^{L_{\mathrm{R}}-1} \widehat{T}^{(s_y)\dagger}_{m\hat{\boldsymbol{y}}}\,\widehat{T}^{(s_x)\dagger}_{n\hat{\boldsymbol{x}}}\,\widehat{O}_{\mu}(\mathbf{0})\,\widehat{T}^{(s_x)}_{n\hat{\boldsymbol{x}}}\,\widehat{T}^{(s_y)}_{m\hat{\boldsymbol{y}}} \\
&= \sum_{m=0}^{L_{\mathrm{R}}-1} \widehat{T}^{(s_y)\dagger}_{m\hat{\boldsymbol{y}}} \left(\widehat{T}^{(s_x)\dagger}_{L_{\mathrm{R}}\hat{\boldsymbol{x}}}\,\widehat{O}_{\mu}(\mathbf{0})\,\widehat{T}^{(s_x)}_{L_{\mathrm{R}}\hat{\boldsymbol{x}}} - \widehat{O}_{\mu}(\mathbf{0}) \right) \widehat{T}^{(s_y)}_{m\hat{\boldsymbol{y}}} = 0
\end{aligned} \tag{C.17}$$

where, after an evident cancellation of most of the terms, the last step follows trivially if $s_x = +1$, since $\widehat{T}^{(+)}_{L_{\mathrm{R}}\hat{\boldsymbol{x}}} = 1$, while it follows from the fact that $\widehat{T}^{(-)}_{L_{\mathrm{R}}\hat{\boldsymbol{x}}} = \widehat{\mathcal{P}}$ and the assumption that $\widehat{\mathcal{P}}^{\dagger}\widehat{O}_{\mu}\widehat{\mathcal{P}} = \widehat{O}_{\mu}$ if $s_x = -1$. A similar derivation applies to $\widehat{T}^{(s_y)}_{\hat{\boldsymbol{y}}}$.

Appendix D

Effective Hamiltonian for digital evolution

Let us consider, to start, the TFIM, and let us analyze the adiabaticity of the associated digital evolution operator $\widehat{U}_m = \widehat{U}^{\mathrm{TFIM}}(\boldsymbol{\gamma}_m)$. To make the analogy with the continuous-time evolution stronger, we introduce an effective Hamiltonian $\widehat{H}_m^{\mathrm{eff}}$ and a time discretization $\Delta t_m > 0$ which satisfy

$$\mathrm{e}^{-i\frac{\Delta t_m}{\hbar}\widehat{H}_m^{\mathrm{eff}}} \stackrel{\mathrm{def}}{=} \widehat{U}_m = \mathrm{e}^{-i\gamma_{1,m}\widehat{H}_x}\mathrm{e}^{-i\gamma_{2,m}\widehat{H}_{zz}} , \tag{D.1}$$

where an additional condition on the spectrum of $\widehat{H}_m^{\mathrm{eff}}$ must be imposed to ensure an unambiguous definition of the logarithm (*e.g.*, spectrum bounded in $[-\frac{\pi\hbar}{\Delta t_m}, \frac{\pi\hbar}{\Delta t_m}]$). Clearly, the definition given in Eq. (D.1) is closely related to the lowest-order Trotter decomposition. Indeed, for $h = 0$, under the assumption $\gamma_{1,m}, \gamma_{2,m} \ll J^{-1}$ and $\gamma_{1,m} + \gamma_{2,m} > 0$, we can use

$$\begin{aligned} \Delta t_m &= \hbar(\gamma_{2,m} + \gamma_{1,m}) > 0 \quad \text{and} \quad s_m = \frac{\gamma_{2,m}}{\gamma_{2,m} + \gamma_{1,m}} \\ \widehat{H}_m^{\mathrm{eff}} &= s_m\, \widehat{H}_{zz} + (1 - s_m)\, \widehat{H}_x + \mathcal{O}((\Delta t_m)^2) \end{aligned} \tag{D.2}$$

to approximately describe the discrete dynamics of the system. Unfortunately, this assumption does not hold for a generic digital evolution and an indiscriminate application of Eq. (D.2) may lead to incorrect results. In particular, the regular schedule is such that $\gamma_{2,m}^{\mathrm{reg}}, \gamma_{1,m}^{\mathrm{reg}} \approx J^{-1}$ for most values of m, so that Eq. (D.2) cannot be used to get $\widehat{H}_m^{\mathrm{eff}}$. One must then use other methods to compute it.

Although in most cases, computing the exact expression of $\widehat{H}_m^{\mathrm{eff}}$ is extremely complicated, the Jordan-Wigner pseudo-spin description allows us to derive an

exact expression for $\widehat{H}_m^{\rm eff}$ in the ordered Ising/XY chain case. In the pseudo-spin picture each k-vector Hilbert space evolves independently with an effective Hamiltonian given by

$$\begin{aligned}
\widehat{H}_m^{(k)} &= \frac{i\hbar}{\Delta t_m}\log\left[{\rm e}^{-i\gamma_{1,m}\hat{\mathcal{H}}_x^{(k)}}\,{\rm e}^{-i\gamma_{2,m}\hat{\mathcal{H}}_z^{(k)}}\right] \\
&= \frac{i\hbar}{\Delta t_m}\log\Big[(\cos(2\gamma_{1,m})+i\sin(2\gamma_{1,m})\;\hat{\boldsymbol{z}}\cdot\hat{\boldsymbol{\tau}}_k)(\cos(2\gamma_{2,m})+i\sin(2\gamma_{2,m})\;\hat{\boldsymbol{b}}_k\cdot\hat{\boldsymbol{\tau}}_k)\Big] \\
&= \frac{i\hbar}{\Delta t_m}\log\left[\cos(\omega_m^{(k)}\Delta t_m)+i\sin(\omega_m^{(k)}\Delta t_m)\;\hat{\boldsymbol{\omega}}_m^{(k)}\cdot\hat{\boldsymbol{\tau}}_k\right] \\
&= \frac{i\hbar}{\Delta t_m}\log\left[{\rm e}^{i\Delta t_m\boldsymbol{\omega}_m^{(k)}\cdot\hat{\boldsymbol{\tau}}_k}\right] = -\hbar\boldsymbol{\omega}_m^{(k)}\cdot\hat{\boldsymbol{\tau}}_k\;, \qquad \text{(D.3)}
\end{aligned}$$

where we used standard properties of the Pauli matrices, $(\mathbf{u}\cdot\hat{\boldsymbol{\tau}})(\mathbf{v}\cdot\hat{\boldsymbol{\tau}}) = \mathbf{u}\cdot\mathbf{v} + i(\mathbf{u}\times\mathbf{v})\cdot\hat{\boldsymbol{\tau}}$ for any two three-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$, and we introduced an effective "magnetic field" $\boldsymbol{\omega}_m^{(k)}$

$$\begin{aligned}
\boldsymbol{\omega}_m^{(k)}\Delta t_m &= \cos(2\gamma_{1,m})\sin(2\gamma_{2,m})\;\hat{\boldsymbol{b}}_k + \cos(2\gamma_{2,m})\sin(2\gamma_{1,m})\;\hat{\boldsymbol{z}} \\
&\qquad - \sin(2\gamma_{1,m})\sin(2\gamma_{2,m})\;\hat{\boldsymbol{z}}\times\hat{\boldsymbol{b}}_k \qquad \text{(D.4)}
\end{aligned}$$

with associated unit vector $\hat{\boldsymbol{\omega}}_m^{(k)} = \boldsymbol{\omega}_m^{(k)}/|\boldsymbol{\omega}_m^{(k)}|$ and the frequency $\omega_m^{(k)} = |\boldsymbol{\omega}_m^{(k)}|$, which can also be shown to be such that:

$$\cos(\omega_m^{(k)}\Delta t_m) = \cos(2\gamma_{1,m})\cos(2\gamma_{2,m}) - \sin(2\gamma_{1,m})\sin(2\gamma_{2,m})\;\hat{\boldsymbol{b}}_k\cdot\hat{\boldsymbol{z}}\;. \qquad \text{(D.5)}$$

We observe that the ambiguity in the logarithm has been transferred to the trigonometric functions.

To address the issue of the "criticality", as seen from the perspective of a digital dynamics, we now look for points in parameter space where the "effective magnetic field" $\boldsymbol{\omega}_m^{(k)}$ *vanishes*. Using the fact that $\hat{\boldsymbol{b}}_k = (-\sin k, 0, \cos k)^T$, after rather simple algebra one can show that:

$$\begin{aligned}
\left|\boldsymbol{\omega}_m^{(k)}\Delta t_m\right|^2 &= \sin^2 2(\gamma_{1,m}-\gamma_{2,m}) + (1-\cos^2 k)\sin^2(2\gamma_{1,m})\sin^2(2\gamma_{2,m}) \\
&\qquad + \frac{(1+\cos k)}{2}\sin(4\gamma_{1,m})\sin(4\gamma_{2,m})\;, \qquad \text{(D.6)}
\end{aligned}$$

where one should recall that the parameters $\gamma_{1,m}$ and $\gamma_{2,m}$ can always be taken to be in the interval $[0, \frac{\pi}{2}]$. One can show that the only k for which such a quantity can possibly vanish is $k = \pi$, and the critical parameters are such that $\gamma_{1,m} \approx \gamma_{2,m}$.

Close to such a point one can expand such a quantity as:

$$\begin{aligned}\left|\boldsymbol{\omega}_m^{(k)}\Delta t_m\right|^2 &= \sin^2 2(\gamma_{1,m}-\gamma_{2,m}) \\ &+ (k-\pi)^2\Big(\sin^2(2\gamma_{1,m})\sin^2(2\gamma_{2,m})+\frac{1}{4}\sin(4\gamma_{1,m})\sin(4\gamma_{2,m})\Big)\end{aligned} \tag{D.7}$$

from which we observe that for $\gamma_{1,m}=\gamma_{2,m}$ the effective field behaves in the usual (linear in $|k-\pi|$) Ising-like fashion:

$$\left|\boldsymbol{\omega}_m^{(k)}\Delta t_m\right| \sim |k-\pi||\sin 2\gamma_{2,m}|+\cdots \tag{D.8}$$

This shows that the "digital criticality" is associated to $\gamma_{1,m}=\gamma_{2,m}$, which translates into $s_m=\frac{1}{2}$ in terms of the s-parameter, precisely as for the continuous-time case.

We now show how to proceed, in principle, with the more general XY chain case, where more unitary operators need to be applied. Consider the following sequence of unitary operators:

$$\mathrm{e}^{-i\gamma_1\hat{\mathcal{H}}_x^{(k)}}\mathrm{e}^{-i\gamma_2\hat{\mathcal{H}}_{yy}^{(k)}}\mathrm{e}^{-i\gamma_3\hat{\mathcal{H}}_x^{(k)}}\mathrm{e}^{-i\gamma_4\hat{\mathcal{H}}_{zz}^{(k)}} = \mathrm{e}^{2i\gamma_1\hat{\boldsymbol{n}}_1\cdot\hat{\boldsymbol{\tau}}_k}\mathrm{e}^{2i\gamma_1\hat{\boldsymbol{n}}_2\cdot\hat{\boldsymbol{\tau}}_k}\mathrm{e}^{2i\gamma_3\hat{\boldsymbol{n}}_3\cdot\hat{\boldsymbol{\tau}}_k}\mathrm{e}^{2i\gamma_4\hat{\boldsymbol{n}}_4\cdot\hat{\boldsymbol{\tau}}_k} \tag{D.9}$$

where we used a convenient notation, introducing the unit vectors $\hat{\boldsymbol{n}}_1=\hat{\boldsymbol{n}}_3=\hat{\boldsymbol{z}}$, $\hat{\boldsymbol{n}}_2=\hat{\boldsymbol{a}}_k$, $\hat{\boldsymbol{n}}_4=\hat{\boldsymbol{b}}_k$. We will now show that there is a very simple iterative procedure which one can set up. Let us start with the first two unitaries (essentially the case we just did for the Ising case). We can write:

$$\begin{aligned}\mathrm{e}^{2i\gamma_1\hat{\boldsymbol{n}}_1\cdot\hat{\boldsymbol{\tau}}_k}\mathrm{e}^{2i\gamma_1\hat{\boldsymbol{n}}_2\cdot\hat{\boldsymbol{\tau}}_k} &= \alpha_{12}+\boldsymbol{\omega}_{12}\cdot\hat{\boldsymbol{\tau}}_k = \cos(2\theta_{12})+i\sin(2\theta_{12})\hat{\boldsymbol{\omega}}_{12}\cdot\hat{\boldsymbol{\tau}}_k \\ &= \mathrm{e}^{2i\theta_{12}\hat{\boldsymbol{\omega}}_{12}\cdot\hat{\boldsymbol{\tau}}_k}\,,\end{aligned} \tag{D.10}$$

where, after easy algebra with Pauli matrices, one finds:

$$\begin{aligned}\alpha_{12} &= \cos(2\gamma_1)\cos(2\gamma_2)-(\hat{\boldsymbol{n}}_1\cdot\hat{\boldsymbol{n}}_2)\sin(2\gamma_1)\sin(2\gamma_2) \\ &\stackrel{\text{def}}{=} \cos(2\theta_{12}) \\ \boldsymbol{\omega}_{12} &= \hat{\boldsymbol{n}}_1\sin(2\gamma_1)\cos(2\gamma_2)+\hat{\boldsymbol{n}}_2\cos(2\gamma_1)\sin(2\gamma_2)-(\hat{\boldsymbol{n}}_1\times\hat{\boldsymbol{n}}_2)\sin(2\gamma_1)\sin(2\gamma_2) \\ &\stackrel{\text{def}}{=} \sin(2\theta_{12})\hat{\boldsymbol{\omega}}_{12}\,.\end{aligned} \tag{D.11}$$

Identical algebra applies to the other two terms:

$$\mathrm{e}^{2i\gamma_3\hat{\boldsymbol{n}}_3\cdot\hat{\boldsymbol{\tau}}_k}\mathrm{e}^{2i\gamma_4\hat{\boldsymbol{n}}_4\cdot\hat{\boldsymbol{\tau}}_k} = \mathrm{e}^{2i\theta_{34}\hat{\boldsymbol{\omega}}_{34}\cdot\hat{\boldsymbol{\tau}}_k}\,, \tag{D.12}$$

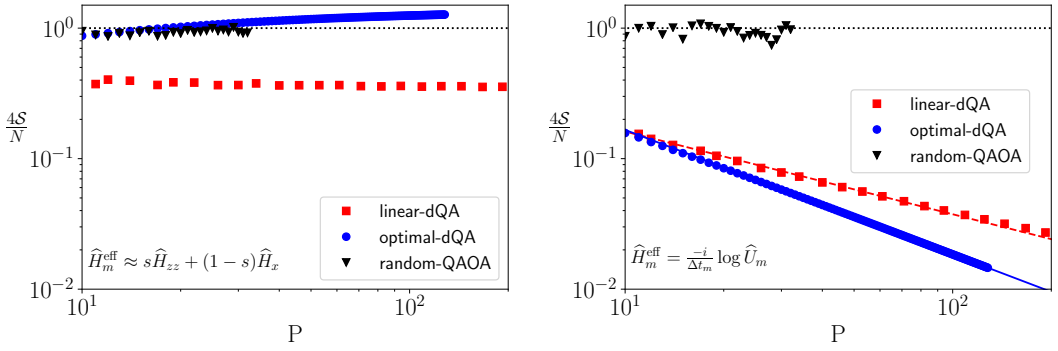

FIGURE D.1: Average Shannon entropy $\mathcal{S}_{\boldsymbol{\gamma}}$ defined in Eq. (4.30) and (5.49) of the main text, for various schedules. The left panel shows the results obtained using the second order approximation in Eq. (D.2). The right panel shows the exact results from Fig. 4.6 of the main text.

and therefore, recursively, also to the product of four terms:

$$\begin{aligned} \mathrm{e}^{2i\gamma_1\hat{\boldsymbol{n}}_1\cdot\hat{\boldsymbol{\tau}}_k}\mathrm{e}^{2i\gamma_1\hat{\boldsymbol{n}}_2\cdot\hat{\boldsymbol{\tau}}_k}\mathrm{e}^{2i\gamma_3\hat{\boldsymbol{n}}_3\cdot\hat{\boldsymbol{\tau}}_k}\mathrm{e}^{2i\gamma_4\hat{\boldsymbol{n}}_4\cdot\hat{\boldsymbol{\tau}}_k} &= \mathrm{e}^{2i\theta_{12}\hat{\boldsymbol{\omega}}_{12}\cdot\hat{\boldsymbol{\tau}}_k}\mathrm{e}^{2i\theta_{34}\hat{\boldsymbol{\omega}}_{34}\cdot\hat{\boldsymbol{\tau}}_k} \\ &= \mathrm{e}^{2i\theta_{1234}\hat{\boldsymbol{\omega}}_{1234}\cdot\hat{\boldsymbol{\tau}}_k}\,, \end{aligned} \tag{D.13}$$

where, clearly, θ_{1234} and $\hat{\boldsymbol{\omega}}_{1234}$ are obtained by iterating the expressions we derived in Eq. (D.11).

This general procedure might be applied to pin down critical points in the parameter space of the digital dynamics of the general XY anisotropic chain. To cut the story short, however, if we need to pin down the combination of parameters which appears at (or close to) $k = \pi$, we just need to notice that the various terms commute, since $\hat{\boldsymbol{a}}_\pi = \hat{\boldsymbol{b}}_\pi = -\hat{\boldsymbol{z}}$, hence, for $k \to \pi$ we would have:

$$\begin{aligned} \mathrm{e}^{-i\gamma_1\hat{\mathcal{H}}_x^{(k\to\pi)}}\mathrm{e}^{-i\gamma_2\hat{\mathcal{H}}_{yy}^{(k\to\pi)}}\mathrm{e}^{-i\gamma_3\hat{\mathcal{H}}_x^{(k\to\pi)}}\mathrm{e}^{-i\gamma_4\hat{\mathcal{H}}_{zz}^{(k\to\pi)}} &= \mathrm{e}^{2i\gamma_1\hat{\tau}^z}\mathrm{e}^{-2i\gamma_2\hat{\tau}^z}\mathrm{e}^{2i\gamma_3\hat{\tau}^z}\mathrm{e}^{-2i\gamma_4\hat{\tau}^z} \\ &= \mathrm{e}^{-2i(\gamma_4-\gamma_3+\gamma_2-\gamma_1)\hat{\tau}^z}\,. \end{aligned} \tag{D.14}$$

This shows why the combination $\gamma_4 - \gamma_3 + \gamma_2 - \gamma_1$ plays a special role, and nearly vanishes in the central part of the optimal regular schedules.

In Fig. D.1 we show that the average Shannon entropy $\mathcal{S}$, defined in Eqs. (4.30) and (5.49) of the main text, when computed using the approximate effective Hamiltonian given in Eq. (D.2) does not correctly signal the adiabaticity of *optimal*-dQA schedule.

Appendix E

Residual energy bounds for the one-dimensional XY model

E.1 Variational Bound

We consider the target Hamiltonian for the one-dimensional XY model:

$$\begin{aligned}\widehat{H}_{\rm XY} &= h\widehat{H}_x + \frac{1-\delta}{2}\widehat{H}_{yy} + \frac{1+\delta}{2}\widehat{H}_{zz} \\ &= -h\sum_{j=1}^{L}\hat{\sigma}_j^x + \frac{1-\delta}{2}\sum_{j=1}^{N}\hat{\sigma}_j^y\hat{\sigma}_{j+1}^y + \frac{1+\delta}{2}\sum_{j=1}^{N}\hat{\sigma}_j^z\hat{\sigma}_{j+1}^z\,. \end{aligned} \tag{E.1}$$

The corresponding reduced Hamiltonian (defined on the reduced chain of an even length $L_{\rm R}$) is

$$\begin{aligned}\hat{\mathcal{H}}^{(\pm)} &= h\hat{\mathcal{H}}_x^{(\pm)} + \frac{1-\delta}{2}\hat{\mathcal{H}}_{yy}^{(\pm)} + \frac{1+\delta}{2}\hat{\mathcal{H}}_{zz}^{(\pm)} \\ &= h(-1\pm 1) + \sum_{k}^{\mathcal{K}^{(\pm)}}\hat{\mathcal{H}}^{(k)}\,, \end{aligned} \tag{E.2}$$

where we have introduced

$$\begin{aligned}\hat{\mathcal{H}}^{(k)} &= h\hat{\mathcal{H}}_x^{(k)} + \frac{1-\delta}{2}\hat{\mathcal{H}}_{yy}^{(k)} + \frac{1+\delta}{2}\hat{\mathcal{H}}_{zz}^{(k)} \\ &= -2\left(h\hat{\boldsymbol{z}} + \frac{1-\delta}{2}\hat{\boldsymbol{a}}_k + \frac{1+\delta}{2}\hat{\boldsymbol{b}}_k\right)\cdot\hat{\boldsymbol{\tau}}_k = -\boldsymbol{\omega}_k(h,\delta)\cdot\frac{\hat{\boldsymbol{\tau}}}{2}\,, \end{aligned} \tag{E.3}$$

and defined the vector $\boldsymbol{\omega}_k(h,\delta) = 4(-\delta\sin k, 0, h+\cos k)^T$.

The ground state energy density of the reduce chain is therefore:

$$\varepsilon_{\mathrm{R}}^{(\pm)}(\mathbf{g}) = (-1 \pm 1)\frac{h}{L_{\mathrm{R}}} - \frac{1}{2L_{\mathrm{R}}}\sum_{k}^{\mathcal{K}^{(\pm)}} \left|\boldsymbol{\omega}_k(h,\delta)\right| , \tag{E.4}$$

where $\mathbf{g} = (h, \delta)$. For a reduced system with ABC this leads to following expression for the energy density:

$$\varepsilon_{\mathrm{R}}^{(-)}(\mathbf{g}) \;=\; -2\frac{h}{L_{\mathrm{R}}} - \frac{2}{L_{\mathrm{R}}}\sum_{n=1}^{\frac{L_{\mathrm{R}}}{2}-1}\sqrt{\delta^2\sin^2\frac{2\pi n}{L_{\mathrm{R}}} + \left(h + \cos\frac{2\pi n}{L_{\mathrm{R}}}\right)^2} , \tag{E.5}$$

while for the full system (with PBC) the energy density is

$$\begin{aligned} e_{\mathrm{gs}}(\mathbf{g}) &= -e_{\max}(\mathbf{g}) \\ &= -\frac{2}{L_{\mathrm{R}}}\sum_{n=1}^{\frac{L}{2}}\sqrt{\delta^2\sin^2\frac{(2n-1)\pi}{2L} + \left(h + \cos\frac{(2n-1)\pi}{2L}\right)^2} . \end{aligned} \tag{E.6}$$

When considering the XY model, these expressions provide us with an analytical expression for the variation bound for the residual energy derived in the main text:

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\boldsymbol{\gamma};\mathbf{g}_{\mathrm{T}}) \geq \epsilon_{\mathrm{P}}^{\mathrm{var}}(\mathbf{g}_{\mathrm{T}}) = \frac{\varepsilon_{\mathrm{R}}^{(-)}(\mathbf{g}_{\mathrm{T}}) - e_{\mathrm{gs}}(\mathbf{g}_{\mathrm{T}})}{e_{\max}(\mathbf{g}_{\mathrm{T}}) - e_{\mathrm{gs}}(\mathbf{g}_{\mathrm{T}})} . \tag{E.7}$$

We now proceed in deriving the asymptotic expansion of the bound for large L_{R}. We can rewrite $\varepsilon_{\mathrm{R}}(\mathbf{g})$ as

$$\begin{aligned} \varepsilon_{\mathrm{R}}^{(-)}(\mathbf{g}) &= -2\frac{h}{L_{\mathrm{R}}} - \frac{1}{2L_{\mathrm{R}}}\sum_{n=1}^{\frac{L_{\mathrm{R}}}{2}-1}\left|\boldsymbol{\omega}_{k_n}\right| \\ &= -2\frac{h}{L_{\mathrm{R}}} + \frac{1}{2L_{\mathrm{R}}}\left(\left|\boldsymbol{\omega}_{k=0}\right| + \left|\boldsymbol{\omega}_{k=\pi}\right|\right) - \frac{1}{2L_{\mathrm{R}}}\sum_{n=0}^{\frac{L_{\mathrm{R}}}{2}}\left|\boldsymbol{\omega}_{k_n}\right| \end{aligned} \tag{E.8}$$

Next we transform the sum over n into an integral over k by using the Euler-MacLaurin formula:

$$\begin{aligned} \sum_{a<n\leq b} f(n) &= \int_a^b \mathrm{d}x\, f(x) + \sum_{m=1}^{\mathrm{M}}\frac{(-1)^m B_m}{m!}\Big(f^{(m-1)}(b) - f^{(m-1)}(a)\Big) \\ &\quad - \int_a^b \mathrm{d}x \frac{B_{\mathrm{M}}(\{1-x\})}{\mathrm{M}!} f^{(\mathrm{M})}(x) , \end{aligned} \tag{E.9}$$

where $a < b \in \mathbb{Z}$, $B_m = B_m(x=0)$ are Bernoulli numbers, $B_m(x)$ are Bernoulli

polynomials and the last (Poisson's remainder) term involves the fractional part $\{x\} = x - \lfloor x \rfloor$ of a real number. Stopping to second order (M = 2), using $B_1 = -1/2$ and $B_2 = 1/6$, and neglecting the Poisson's remainder, we get:

$$\frac{1}{2L_{\mathrm{R}}} \sum_{n=0}^{\frac{L_{\mathrm{R}}}{2}} \left|\boldsymbol{\omega}_{k_n}\right| = \int_0^{\pi} \frac{\mathrm{d}k}{4\pi} \left|\boldsymbol{\omega}_k\right| + \frac{1}{4L_{\mathrm{R}}} \Big(\left|\boldsymbol{\omega}_{\pi}\right| + \left|\boldsymbol{\omega}_0\right| \Big) + \frac{\pi}{12L_{\mathrm{R}}^2} \frac{\mathrm{d}\left|\boldsymbol{\omega}_k\right|}{\mathrm{d}k} \Bigg|_{k=0}^{k=\pi} + \mathcal{O}(L_{\mathrm{R}}^{-3}) \,. \tag{E.10}$$

Using Eq. (E.10) into the expression for $\varepsilon_{\mathrm{R}}^{(-)}$ we get

$$\varepsilon_{\mathrm{R}}^{(-)}(\mathbf{g}) = e_{\mathrm{gs}}^{(\infty)}(\mathbf{g}) - \frac{2h - |h+1| - |h-1|}{L_{\mathrm{R}}} - \frac{\pi}{12L_{\mathrm{R}}^2} \frac{\mathrm{d}\left|\boldsymbol{\omega}_k\right|}{\mathrm{d}k} \Bigg|_{k=0}^{k=\pi} + \mathcal{O}(L_{\mathrm{R}}^{-3}) \,. \tag{E.11}$$

Here we have introduced the ground-state energy density for the infinite chain

$$e_{\mathrm{gs}}^{(\infty)}(h,\delta) = -\int_0^{\pi} \frac{\mathrm{d}k}{4\pi} \left|\boldsymbol{\omega}_k(h,\delta)\right| \,, \tag{E.12}$$

which can be expressed in terms of elliptic integrals Maciążek and Wojtkiewicz [2016]. Notice also the derivative term which does not contribute unless $|\boldsymbol{\omega}_k|$ has a singularity at the border of the interval, which happens to occur for $h = \pm 1$. Similar considerations apply to the higher-order derivatives, which generally cancel when the function is smooth enough.

The expansion of $\varepsilon_{\mathrm{R}}^{(-)}(\mathbf{g})$ can be specified for the cases of different intervals of h obtaining

$$\varepsilon_{\mathrm{R}}^{(-)}(\mathbf{g}) = e_{\mathrm{gs}}^{(\infty)}(h,\delta) + \begin{cases} \frac{-4h}{L_{\mathrm{R}}} + \mathcal{O}(L_{\mathrm{R}}^{-3}) & \text{if } h < -1 \\ \frac{2-2h}{L_{\mathrm{R}}} + \mathcal{O}(L_{\mathrm{R}}^{-3}) & \text{if } -1 \leq h < 1 \\ \frac{4\pi|\delta|}{3L_{\mathrm{R}}^2} + \mathcal{O}(L_{\mathrm{R}}^{-3}) & \text{if } h = 1 \\ \mathcal{O}(L_{\mathrm{R}}^{-3}) & \text{if } h > 1 \end{cases} . \tag{E.13}$$

The variational bound can be written as

$$\epsilon^{\text{var}}(\mathbf{g}_{\text{T}}) = \frac{1}{2\left|e_{\text{gs}}^{(\infty)}(h,\delta)\right|} \times \begin{cases} \frac{-4h}{L_{\text{R}}} + \mathcal{O}(L_{\text{R}}^{-3}) & \text{if } h < -1 \\ \frac{2-2h}{L_{\text{R}}} + \mathcal{O}(L_{\text{R}}^{-3}) & \text{if } -1 \leq h < 1 \\ \frac{4\pi|\delta|}{3L_{\text{R}}^2} + \mathcal{O}(L_{\text{R}}^{-3}) & \text{if } h = 1 \\ \mathcal{O}(L_{\text{R}}^{-3}) & \text{if } h > 1 \end{cases} . \tag{E.14}$$

Notice how the scaling changes when crossing $h = 1$.

The computation just performed does not capture higher orders. However, the smoothness of $|\boldsymbol{\omega}_k|$ for $h > 1$ leads to a cancellation of all corrections, and the convergence occurs exponentially fast.

E.2 Tightness of the bound for $\mathrm{P} = 1$

We consider the case where only two driving Hamiltonians $\widehat{H}_x$ and $\widehat{H}_{zz}$ are used. For $\mathrm{P} = 1$ and $N \geq 2\mathrm{P}+2 = 4$ the variational bound can be attained exactly for all targets. As we already mentioned the exact expression of the residual energy can be equivalently computed using a reduced chain with ABC. Specifying $L_{\text{R}} = 4$ in Eq. (A.12) we get that there is a single relevant positive k-mode $\tilde{\mathcal{K}}^{(-)} = \{\frac{\pi}{2}\}$. The Hamiltonians describing the single mode $k = \frac{\pi}{2}$ can be read form Eqs. (A.24)—(A.26), giving

$$\hat{\mathcal{H}}_x^{(\frac{\pi}{2})} = -2\hat{\tau}^z_{\frac{\pi}{2}} \tag{E.15}$$

$$\hat{\mathcal{H}}_{yy}^{(\frac{\pi}{2})} = -2\,\hat{\tau}^x_{\frac{\pi}{2}} \tag{E.16}$$

$$\hat{\mathcal{H}}_{zz}^{(\frac{\pi}{2})} = +2\,\hat{\tau}^x_{\frac{\pi}{2}} \,. \tag{E.17}$$

It is then straightforward to compute the energy density $e_{\mathrm{P}=1}(\boldsymbol{\gamma};\mathbf{g}_{\text{T}})$ of the final state:

$$\begin{aligned} e_{\mathrm{P}=1}(\boldsymbol{\gamma};\mathbf{g}_{\text{T}}) &= -\frac{h}{2} + \frac{1}{4}\langle\uparrow|\mathrm{e}^{i\gamma_{2,1}\hat{\mathcal{H}}_{zz}^{(\frac{\pi}{2})}}\mathrm{e}^{i\gamma_{1,1}\hat{\mathcal{H}}_x^{(\frac{\pi}{2})}}\Big(h\hat{\mathcal{H}}_x^{(\frac{\pi}{2})} + \frac{1-\delta}{2}\hat{\mathcal{H}}_{yy}^{(\frac{\pi}{2})} \\ &\qquad + \frac{1+\delta}{2}\hat{\mathcal{H}}_{yy}^{(\frac{\pi}{2})}\Big)\mathrm{e}^{-i\gamma_{1,1}\hat{\mathcal{H}}_x^{(\frac{\pi}{2})}}\mathrm{e}^{-i\gamma_{2,1}\hat{\mathcal{H}}_{zz}^{(\frac{\pi}{2})}}|\uparrow\rangle \\ &= -\frac{h}{2} + \frac{1}{2}\langle\uparrow|\mathrm{e}^{i2\gamma_{2,1}\hat{\tau}^x_{\frac{\pi}{2}}}\mathrm{e}^{-i2\gamma_{1,1}\hat{\tau}^z_{\frac{\pi}{2}}}\left(\delta\hat{\tau}^x_{\frac{\pi}{2}} - h\hat{\tau}^z_{\frac{\pi}{2}}\right)\mathrm{e}^{i2\gamma_{1,1}\hat{\tau}^z_{\frac{\pi}{2}}}\mathrm{e}^{-i2\gamma_{2,1}\hat{\tau}^x_{\frac{\pi}{2}}}|\uparrow\rangle \\ &= -\frac{h}{2} - \frac{1}{2}\Big(\delta\sin(4\gamma_{2,1})\sin(4\gamma_{1,1}) + h\cos(4\gamma_{2,1})\Big) \,. \end{aligned} \tag{E.18}$$

The minimum value is obtained for

$$(\gamma^*_{1,1}, \gamma^*_{2,1}) \;=\; \pm\left(\frac{\pi}{8}, \frac{1}{4}\arccos\frac{h}{\sqrt{h^2+\delta^2}}\right)\,, \tag{E.19}$$

where the corresponding optimal value attained by the residual energy is

$$\begin{aligned}\min_{\boldsymbol{\gamma}}\left[e_{\mathrm{P}=1}(\boldsymbol{\gamma};\mathbf{g}_{\mathrm{T}})\right] &= -\frac{h}{2} - \frac{1}{2}\Big(\delta\sin(4\gamma^*_{2,1})\sin(4\gamma^*_{1,1}) + h\cos(4\gamma^*_{2,1})\Big)\\ &= -\frac{h}{2} - \frac{\sqrt{\delta^2+h^2}}{2} = \varepsilon^{(-)}_{\mathrm{R}}(\mathbf{g}_{\mathrm{T}})\,.\end{aligned} \tag{E.20}$$

Here we used the expression of $\varepsilon^{(-)}_{\mathrm{R}}(\mathbf{g}_{\mathrm{T}})$ given in Eq. (E.5) (with $L_{\mathrm{R}} = 4$). This proves that for $\mathrm{P} = 1$, the variational bound is tight in the whole phase diagram of the XY model.

E.3 Tightness of the bound for $\mathrm{P} > 1$

We still consider the case where only two driving Hamiltonians $\widehat{H}_x$ and $\widehat{H}_{zz}$ are used. Already for $\mathrm{P} = 2, \cdots$ and $N > 2\mathrm{P}$ the exact expression for the residual energy is more involved, and finding an analytical expression for the optimal values is difficult. However, for small values of P (more precisely, for $\mathrm{P} = 2, \cdots 8$), since the expectation value of the energy depends only on 2P variational parameters, it still possible to carry out a numerical *global optimization* in the space $\gamma \in [0, \frac{\pi}{2}]^{2\mathrm{P}}$. In Fig. E.1(a) we show the results of the optimization for $h = 0$ as a function of δ. For fixed P, the relative discrepancy between the optimal value and the variational bound vanishes when δ is above a threshold $\overline{\delta}_{\mathrm{P}}$. In particular for $\delta > \overline{\delta}_{\mathrm{P}}$ we find that the $\epsilon^{\mathrm{res}}_{\mathrm{P}}(\boldsymbol{\gamma}^{\mathrm{opt}}, \mathbf{g}_{\mathrm{T}}) = \epsilon^{\mathrm{var}}_{\mathrm{P}}(\mathbf{g}_{\mathrm{T}})$ up to machine precision (10^{-14}). In Fig. E.1(b) we repeat the same analysis over the whole XY model phase diagram and show that there are small (colored) regions close to the line $\delta = 0$, where the variational bound is not exactly attained through optimization of the parameters. Outside of such region, however, the variational bound can be attained exactly, and there are 2^{P} degenerate global minima (we have checked this numerically for small values P: we have no analytical proof for that). The size of the region where $\epsilon^{\mathrm{var}}_{\mathrm{P}} \neq \epsilon^{\mathrm{res}}_{\mathrm{P}}$ shrinks when P is increased. Figure E.1(c) shows that when targeting ($|h| < 1, \delta = 0$), although for any finite P the variational bound cannot be exactly attained, the discrepancy goes to zero for $\mathrm{P} \to \infty$.

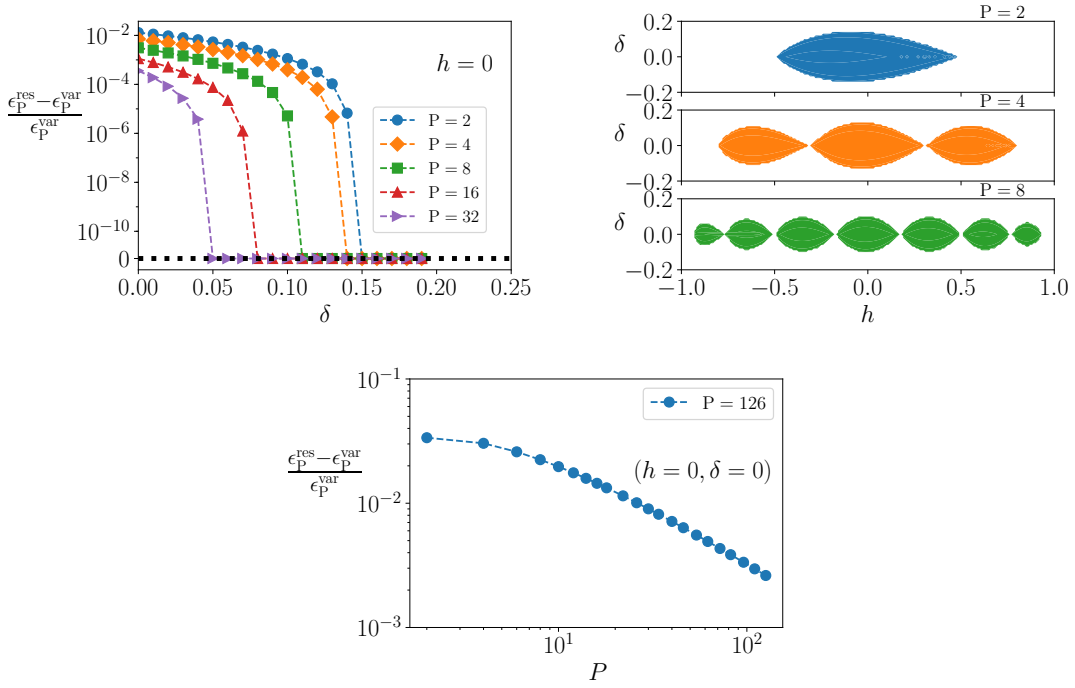

FIGURE E.1: Tightness of the variational bound for ϵ_P^{res} in the XY model. (a) Relative distance between the optimized residual energy ϵ_P^{res} and the exact variational bound ϵ_P^{var} for $P = 2, 4, 8, 16, 32$. The y axes is plotted in a symmetric logarithm scale (linear for $y < 10^{-10}$ and logarithmic for $y > 10^{-10}$). The black dotted line corresponds to case where the variational bound is tight $\epsilon_P^{var} = \epsilon_P^{res}$. (b) Regions in the XY phase diagram where the variational bound is not tight $\epsilon_P^{var} \neq \epsilon_P^{res}$ for $P = 2$. (c) Relative distance between the optimized residual energy ϵ_P^{res} and the exact variational bound ϵ_P^{var} as for the target $(h = 0, \delta = 0)$ as a function of P

.

Bibliography

D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, *Adiabatic quantum computation is equivalent to standard quantum computation,* SIAM Review **50**, 755 (2008).

T. Albash and D. A. Lidar, *Adiabatic quantum computation,* Rev. Mod. Phys. **90**, 015002 (2018).

A. Ambainis, *On physical problems that are slightly more difficult than QMA,* arXiv:1312.4758 [quant-ph] (2013).

L. Arceci, S. Barbarino, D. Rossini, and G. E. Santoro, *Optimal working point in dissipative quantum annealing,* Phys. Rev. B **98**, 064307 (2018).

C. Baldassi and R. Zecchina, *Efficiency of quantum vs. classical annealing in non-convex learning problems,* Proceedings of the National Academy of Sciences **115**, 1457 (2018).

V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, *The quantum adiabatic algorithm applied to random optimization problems: The quantum spin glass perspective,* Physics Reports **523**, 127 (2013).

R. Barankov and A. Polkovnikov, *Optimal nonlinear passage through a quantum critical point,* Phys. Rev. Lett. **101**, 076801 (2008).

R. Barends, A. Shabani, L. Lamata, J. Kelly, A. Mezzacapo, U. L. Heras, R. Babbush, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, E. Solano, H. Neven, and J. M. Martinis, *Digitized adiabatic quantum computing with a superconducting circuit,* Nature **534**, 222 (2016).

T. Barthel and M. Kliesch, *Quasilocality and efficient simulation of markovian quantum dynamics,* Phys. Rev. Lett. **108**, 230504 (2012).

F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, 2017).

C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, *Strengths and weaknesses of quantum computing,* SIAM Journal on Computing **26**, 1510 (1997).

H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, *et al.*, *Probing many-body dynamics on a 51-atom quantum simulator,* Nature **551**, 579 (2017).

I. Bloch, J. Dalibard, and S. Nascimbene, *Quantum simulations with ultracold quantum gases,* Nature Physics **8**, 267 (2012).

S. Boixo, T. F. Rønnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, *Evidence for quantum annealing with more than one hundred qubits,* Nat Phys **10**, 218 (2014).

E. Boros and P. L. Hammer, *The max-cut problem and quadratic 0–1 optimization; polyhedral aspects, relaxations and bounds,* Annals of Operations Research **33**, 151 (1991).

E. Boros and P. L. Hammer, *Pseudo-boolean optimization,* Discrete Applied Mathematics **123**, 155 (2002).

S. B. Bravyi and A. Y. Kitaev, *Quantum codes on a lattice with boundary,* arXiv:9811052 [quant-ph] (1998).

S. Bravyi, D. P. Divincenzo, R. Oliveira, and B. M. Terhal, *The complexity of stoquastic local hamiltonian problems,* Quantum Info. Comput. **8**, 361 (2008).

C. Brif, R. Chakrabarti, and H. Rabitz, *Control of quantum phenomena: past, present and future,* New Journal of Physics **12**, 075008 (2010).

J. Brooke, D. Bitko, F. T. Rosenbaum, and G. Aeppli, *Quantum annealing of a disordered magnet,* Science **284**, 779 (1999).

S. P. Brooks and G. O. Roberts, *Convergence assessment techniques for markov chain monte carlo,* Statistics and Computing **8**, 319 (1998).

L. Brualla, K. Sakkos, J. Boronat, and J. Casulleras, *Higher order and infinite trotter-number extrapolations in path integral monte carlo,* The Journal of Chemical Physics **121**, 636 (2004).

M. Bukov, A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, *Reinforcement learning in different phases of quantum control,* Phys. Rev. X **8**, 031086 (2018).

M. Calandra Buonaura and S. Sorella, *Numerical study of the two-dimensional heisenberg model using a green function monte carlo technique with a fixed number of walkers,* Phys. Rev. B **57**, 11446 (1998).

T. Caneva, R. Fazio, and G. E. Santoro, *Adiabatic quantum dynamics of a random ising chain across its quantum critical point,* Phys. Rev. B **76**, 144427 (2007).

L. Capriotti, F. Becca, A. Parola, and S. Sorella, *Resonating valence bond wave functions for strongly frustrated spin systems,* Phys. Rev. Lett. **87**, 097201 (2001).

D. M. Ceperley, *Path integrals in the theory of condensed helium,* Rev. Mod. Phys. **67**, 279 (1995).

K. Choo, T. Neupert, and G. Carleo, *Two-dimensional frustrated J_1-J_2 model studied with neural network quantum states,* Phys. Rev. B **100**, 125124 (2019).

P. W. Claeys, M. Pandey, D. Sels, and A. Polkovnikov, *Floquet-engineering counterdiabatic protocols in quantum many-body systems,* Phys. Rev. Lett. **123**, 090602 (2019).

G. E. Crooks, *Performance of the Quantum Approximate Optimization Algorithm on the Maximum Cut Problem,* arXiv:1811.08419 [quant-ph] (2018).

E. Crosson and A. W. Harrow, in *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)* (2016) pp. 714–723.

T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, *Undecidability of the spectral gap,* Nature **528**, 207 (2015).

D. D'Alessandro, *Introduction to quantum control and dynamics* (Chapman and Hall/CRC, 2007).

C. De Grandi and A. Polkovnikov, Adiabatic perturbation theory: From landau–zener problem to quenching through a quantum critical point, in *Quantum Quenching, Annealing and Computation*, edited by A. K. Chandra, A. Das, and B. K. Chakrabarti (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010) pp. 75–114.

V. S. Denchev, S. Boixo, S. V. Isaksov, N. Ding, R. Babbush, V. Smelyanskiy, and J. Martinis, *What is the computational value of finite-range tunneling?* Phys. Rev. X **6**, 031015 (2016).

A. Dranov, J. Kellendonk, and R. Seiler, *Discrete time adiabatic theorems for quantum mechanical systems,* Journal of Mathematical Physics **39**, 1340 (1998).

J. Dziarmaga, *Dynamics of a quantum phase transition: Exact solution of the quantum ising model,* Phys. Rev. Lett. **95**, 245701 (2005).

J. Dziarmaga, *Dynamics of a quantum phase transition in the random ising model: Logarithmic dependence of the defect density on the transition rate,* Phys. Rev. B **74**, 064416 (2006).

E. Farhi and S. Gutmann, *Analog analogue of a digital quantum computation,* Phys. Rev. A **57**, 2403 (1998).

E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, *A quantum adiabatic evolution algorithm applied to random instances of an NP-Complete problem,* Science **292**, 472 (2001).

E. Farhi, J. Goldstone, and S. Gutmann, *A Quantum Approximate Optimization Algorithm,* arXiv:1411.4028 [quant-ph] (2014).

R. P. Feynman, *Simulating physics with computers,* International journal of theoretical physics **21**, 467 (1982).

A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, *Quantum annealing: A new method for minimizing multidimensional functions,* Chem. Phys. Lett. **219**, 343 (1994).

D. S. Fisher, *Critical behavior of random transverse-field ising spin chains,* Phys. Rev. B **51**, 6411 (1995).

A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, *Surface codes: Towards practical large-scale quantum computation,* Phys. Rev. A **86**, 032324 (2012).

M. R. Garey and D. S. Johnson, *Computers and Intractability* (Freeman, 1979).

G. Grosso and G. P. Parravicini, *Solid State Physics* (Academic, San Diego, 2000).

L. K. Grover, *Quantum mechanics helps in seraching for a needle in a haystack,* Phys. Rev. Lett. **79**, 325 (1997).

M. Hastings, *Classical and Quantum Bounded Depth Approximation Algorithms,* arXiv:1905.07047 [quant-ph] (2019).

B. Heim, T. F. Rønnow, S. V. Isakov, and M. Troyer, *Quantum versus classical annealing of ising spin glasses,* Science **348**, 215 (2015).

T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular electronic-structure theory* (John Wiley & Sons, 2014).

I. Hen, *How quantum is the speedup in adiabatic unstructured search?* Quantum Information Processing **18**, 162 (2019).

W. W. Ho and T. H. Hsieh, *Efficient variational simulation of non-trivial quantum states,* SciPost Phys. **6**, 29 (2019).

D. A. Huse and D. S. Fisher, *Residual energies after slow cooling of disordered systems,* Phys. Rev. Lett. **57**, 2203 (1986).

S. V. Isakov, G. Mazzola, V. N. Smelyanskiy, Z. Jiang, S. Boixo, H. Neven, and M. Troyer, *Understanding quantum tunneling through quantum monte carlo simulations,* Phys. Rev. Lett. **117**, 180402 (2016).

P. Jordan and E. Wigner, *Über das Paulische Äquivalenzverbot,* Zeitschrift fur Physik **47**, 631 (1928).

T. Kadowaki and H. Nishimori, *Quantum annealing in the transverse ising model,* Phys. Rev. E **58**, 5355 (1998).

S. Katsura, *Statistical Mechanics of the Anisotropic Linear Heisenberg Model,* Physical Review **127**, 1508 (1962).

B. Keimer, S. A. Kivelson, M. R. Norman, S. Uchida, and J. Zaanen, *From quantum matter to high-temperature superconductivity in copper oxides,* Nature **518**, 179 (2015).

T. W. B. Kibble, *Topology of cosmic domains and strings,* Journal of Physics A: Mathematical and General **9**, 1387 (1976).

T. Kibble, *Some implications of a cosmological phase transition,* Physics Reports **67**, 183 (1980).

S. Kirkpatrick, J. C. D. Gelatt, and M. P. Vecchi, *Optimization by simulated annealing,* Science **220**, 671 (1983).

M. Kliesch, C. Gogolin, and J. Eisert, Lieb-robinson bounds and the simulation of time-evolution of local observables in lattice systems, in *Many-Electron Approaches in Physics, Chemistry and Mathematics: A Multidisciplinary View*, edited by V. Bach and L. Delle Site (Springer International Publishing, Cham, 2014) pp. 301–318.

S. Knysh, *Zero-temperature quantum annealing bottlenecks in the spin-glass phase,* Nature communications **7**, 12370 (2016).

C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. Joshi, P. Jurcevic, C. Muschik, P. Silvi, R. Blatt, C. Roos, *et al.*, *Self-verifying variational quantum simulation of lattice models,* Nature **569**, 355 (2019).

E. H. Lieb and D. W. Robinson, *The finite group velocity of quantum spin systems,* Communications in Mathematical Physics **28**, 251 (1972).

E. Lieb, T. Schultz, and D. Mattis, *Two soluble models of an antiferromagnetic chain,* Ann. Phys. **16**, 407 (1961).

S. Lloyd, *Quantum Approximate Optimization is computationally universal,* arXiv:1904.03292 [quant-ph] (2018).

A. Lucas, *Ising formulations of many np problems,* Frontiers in Physics **2**, 5 (2014).

T. Maciążek and J. Wojtkiewicz, *On the phase diagram of the anisotropic xy chain in transverse magnetic field,* Physica A: Statistical Mechanics and its Applications **441**, 131 (2016).

R. Martoňák, G. E. Santoro, and E. Tosatti, *Quantum annealing by the path-integral monte carlo method: The two-dimensional random ising model,* Phys. Rev. B **66**, 094203 (2002).

G. Mazzola, V. N. Smelyanskiy, and M. Troyer, *Quantum monte carlo tunneling from quantum chemistry to quantum annealing,* Phys. Rev. B **96**, 134305 (2017).

G. B. Mbeng, L. Privitera, L. Arceci, and G. E. Santoro, *Dynamics of simulated quantum annealing in random ising chains,* Physical Review B **99**, 064201 (2019).

G. B. Mbeng, R. Fazio, and G. Santoro, *Quantum Annealing: a journey through Digitalization, Control, and hybrid Quantum Variational schemes,* arXiv:1906.08948 [quant-ph] (2019).

A. Messiah, *Quantum mechanics*, Vol. 2 (North-Holland, Amsterdam, 1962).

A. Mizel, D. A. Lidar, and M. Mitchell, *Simple proof of equivalence between adiabatic quantum computation and the circuit model,* Phys. Rev. Lett. **99**, 070502 (2007).

V. Mukherjee, U. Divakaran, A. Dutta, and D. Sen, *Quenching dynamics of a quantum xy spin-$\frac{1}{2}$ chain in a transverse field,* Phys. Rev. B **76**, 174303 (2007).

B. Nachtergaele and R. Sims, in *New Trends in Mathematical Physics*, edited by V. Sidoravičius (Springer Netherlands, Dordrecht, 2009) pp. 591–614.

M. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

J. Nocedal and S. Wright, *Numerical optimization* (Springer Science & Business Media, 2006).

G. Parisi, M. Mézard, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, 1987).

A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, in *NIPS Autodiff Workshop* (2017).

D. Patanè, A. Silva, L. Amico, R. Fazio, and G. E. Santoro, *Adiabatic dynamics in open quantum critical many-body systems,* Phys. Rev. Lett. **101**, 175701 (2008).

D. Patanè, A. Silva, L. Amico, R. Fazio, and G. E. Santoro, *Adiabatic dynamics of a quantum critical system coupled to an environment: Scaling and kinetic equation approaches,* Phys. Rev. B **80**, 2009 (2009).

A. Perdomo-Ortiz, S. E. Venegas-Andraca, and A. Aspuru-Guzik, *A study of heuristic guesses for adiabatic quantum computation,* Quantum Information Processing **10**, 33 (2011).

A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'brien, *A variational eigenvalue solver on a photonic quantum processor,* Nature communications **5**, 4213 (2014).

H. Pichler, S.-T. Wang, L. Zhou, S. Choi, and M. D. Lukin, *Quantum Optimization for Maximum Independent Set Using Rydberg Atom Arrays,* arXiv:1808.10816 [quant-ph] (2018).

A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, *Nonequilibrium dynamics of closed interacting quantum systems,* Rev. Mod. Phys. **83**, 863 (2011).

A. Polkovnikov, *Universal adiabatic dynamics in the vicinity of a quantum critical point,* Phys. Rev. B **72**, 161201 (2005).

J. Preskill, *Quantum Computing in the NISQ era and beyond,* Quantum **2**, 79 (2018).

M. V. Rakov and M. Weyrauch, *Spin-1/2 XXZ Heisenberg chain in a longitudinal magnetic field,* 1908.11362 (2019).

A. T. Rezakhani, W.-J. Kuo, A. Hamma, D. A. Lidar, and P. Zanardi, *Quantum adiabatic brachistochrone,* Phys. Rev. Lett. **103**, 080502 (2009).

J. Roland and N. J. Cerf, *Quantum search by local adiabatic evolution,* Phys. Rev. A **65**, 042308 (2002).

T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, *Defining and detecting quantum speedup,* Science , 1252319 (2014).

J. J. Sakurai and E. D. Commins, *Modern quantum mechanics, revised edition* (AAPT, 1995).

A. W. Sandvik, *Finite-size scaling of the ground-state parameters of the two-dimensional heisenberg model,* Phys. Rev. B **56**, 11678 (1997).

G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, *Theory of quantum annealing of an ising spin glass,* Science **295**, 2427 (2002).

U. Schollwöck, *The density-matrix renormalization group in the age of matrix product states,* Annals of Physics **326**, 96 (2011).

B. Seoane and H. Nishimori, *Many-body transverse interactions in the quantum annealing of thep-spin ferromagnet,* Journal of Physics A: Mathematical and Theoretical **45**, 435301 (2012).

Y. Susa, Y. Yamashiro, M. Yamamoto, I. Hen, D. A. Lidar, and H. Nishimori, *Quantum annealing of the p-spin model under inhomogeneous transverse field driving,* Phys. Rev. A **98**, 042326 (2018).

M. Suzuki, *Relationship between d-dimensional quantal spin systems and (d+1)-dimensional ising systemsequivalence, critical exponents and systematic approximants of the partition function and spin correlations,* Progress of Theoretical Physics **56**, 1454 (1976).

S. Suzuki, *Cooling dynamics of pure and random ising chains,* J. Stat. Mech. , P03032 (2009).

R. H. Swendsen and J.-S. Wang, *Nonuniversal critical dynamics in monte carlo simulations,* Phys. Rev. Lett. **58**, 86 (1987).

M. Troyer and U.-J. Wiese, *Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations,* Phys. Rev. Lett. **94**, 170201 (2005).

Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, *Quantum approximate optimization algorithm for maxcut: A fermionic view,* Phys. Rev. A **97**, 022304 (2018).

M. M. Wauters, R. Fazio, H. Nishimori, and G. E. Santoro, *Direct comparison of quantum and simulated annealing on a fully connected ising ferromagnet,* Phys. Rev. A **96**, 022326 (2017).

D. Wecker, M. B. Hastings, and M. Troyer, *Training a quantum optimizer,* Physical Review A **94**, 022309 (2016).

T. Weymuth and M. Reiher, *Inverse quantum chemistry: Concepts and strategies for rational compound design,* International Journal of Quantum Chemistry **114**, 823 (2014).

U. Wolff, *Collective monte carlo updating for spin systems,* Phys. Rev. Lett. **62**, 361 (1989).

C. Yang and C. Yang, *One-dimensional chain of anisotropic spin-spin interactions,* Physics Letters **20**, 9 (1966).

Z.-C. Yang, A. Rahmani, A. Shabani, H. Neven, and C. Chamon, *Optimizing variational quantum algorithms using pontryagin's minimum principle,* Phys. Rev. X **7**, 021027 (2017).

A. P. Young and H. Rieger, *Numerical study of the random transverse-field Ising spin chain,* Phys. Rev. B **53**, 8486 (1996).

A. P. Young, *Finite-temperature and dynamical properties of the random transverse-field Ising spin chain,* Phys. Rev. B **56**, 11691 (1997).

M. Yuezhen Niu, S. Lu, and I. L. Chuang, *Optimizing QAOA: Success Probability and Runtime Dependence on Circuit Depth,* arXiv:1905.12134 [quant-ph] (2019).

C. Zalka, *Grover's quantum searching algorithm is optimal,* Phys. Rev. A **60**, 2746 (1999).

T. Zanca and G. E. Santoro, *Quantum annealing speedup over simulated annealing on random ising chains,* Phys. Rev. B **93**, 224431 (2016).

L. Zeng, J. Zhang, and M. Sarovar, *Schedule path optimization for adiabatic quantum computing and optimization,* Journal of Physics A: Mathematical and Theoretical **49**, 165305 (2016).

Q. Zhuang, *Increase of degeneracy improves the performance of the quantum adiabatic algorithm,* Phys. Rev. A **90**, 052317 (2014).

W. H. Zurek, U. Dorner, and P. Zoller, *Dynamics of a quantum phase transition,* Phys. Rev. Lett. **95**, 105701 (2005).

W. H. Zurek, *Cosmological experiments in superfluid helium?* Nature **317**, 505 (1985).

W. Zurek, *Cosmological experiments in condensed matter systems,* Physics Reports **276**, 177 (1996).